MOLECULAR ANALYSIS OF FATTY ACID PEROXYGENASE INVOLVED IN THE BIOSYNTHESIS OF EPOXY FATTY ACIDS IN OATS (*Avena sativa*)

A Thesis Submitted to the

College of Graduate Studies and Research

In Partial Fulfillment of the Requirements

For the Degree of Master of Science

In the Department of Food and Bioproduct Sciences

College of Agriculture and Bioresources

University of Saskatchewan

Saskatoon, Saskatchewan

Canada

By

Indika Gayani Benaragama

2015

PERMISSION TO USE

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

DISCLAIMER

Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

The Head of the Department of Food and Bioproduct Sciences
51 Campus Drive, University of Saskatchewan
Saskatoon, Saskatchewan S7N 5A8
Canada

OR

Dean College of Graduate Studies and Research
University of Saskatchewan
107 Administration Place
Saskatoon, Saskatchewan S7N 5A2 Canada

ABSTRACT

Oat is known to synthesize several epoxy fatty acids in seeds using peroxygenase (PXG), a type of hydroperoxide-dependent epoxygenase. This thesis aims to molecularly clone and functionally characterize the PXG genes from oat developing seeds. The research started with identifying additional PXG genes from oat expressed sequence tag (EST) databases using a previously identified oat peroxygenase AsPXG1 as a query sequence. This resulted in the identification of six homologous contig sequences from the EST data bases. Of them, two contigs with high sequence similarity and alignment with plant PXG/caleosin proteins were selected for cloning and functional analysis. Rapid amplification of cDNA ends (RACE) and reverse transcriptase-polymerase chain reaction (RT-PCR) were used to retrieve the full length cDNAs of the contigs, which resulted in identification of three putative PXG genes, *AsPXG2*, *AsPXG3* and *AsPXG4*. Open reading frame (ORF) of *AsPXG2* is 702 bp long encoding a polypeptide of 233 amino acids, while ORFs of both *AsPXG3* and *AsPXG4* are 627 bp in length coding for 208 amino acids. All these putative peroxygenases comprise a single transmembrane domain, presumably for lipid droplet anchoring, conserved hisditine residues for heme-binding and a conserved EF-hand motif for calcium-binding. To functionally characterize the three genes, their ORFs were individually expressed in *Escherichia coli*/*Pichia pastoris*. The enzymatic assays showed that all transformants produced 9,10-epoxystearic acid methyl ester in the presence of oleic acid methyl ester and cumene hydroperoxide, indicating all three genes encode functional peroxygenase. AsPXG3 has the highest specific activity at 42 μmol/mg/min with about 25% substrate conversion efficiency. Substrate specificity assays on free fatty acids showed that AsPXG3 could epoxidize all mono- and poly-unsaturated fatty acids tested, with linolenic acid (C18:3-9*c*,12*c*,15*c*) being the most preferred substrate. Site-directed mutagenesis of three conserved histidines and nine conserved residues surrounding the histidines in AsPXG3 showed that substitution of the first conserved histidine at position 32 (H1) and the third conserved histidine at position 102 (H3) with alanine respectively resulted in complete loss of the enzymatic activity, while substitution of the second conserved histidine at position 98 (H2) resulted in only slight reduction of the activity, indicating that only H1 and H3 are absolutely essential and probably involved in heme-binding for the peroxygenase. Substitution of leucine at

position 29 (M1), isoleucine at position 97 (M5), and lysine at position 101 (M8) with alanine reduced the enzymatic activity on oleic acid methyl ester by more than 80% relative to the wild type enzyme, indicating these three residues are also very important for catalytic activity. The activity of M1, M5 and M8 mutants was also drastically reduced on all other free mono-unsaturated fatty acids tested (>60%). However, to linolenic acid, M5 showed only slight reduction of the activity (~15%) and M8 even increased the activity by 12% relative to the wild type enzyme. These results suggest that these conserved residues might play roles in defining the shape and size of the catalytic site for interaction of the heme with fatty acid substrates.

ACKNOWLEDGEMENT

I would like to take this opportunity to express my sincere gratitude to my supervisor Dr. Xiao Qiu for his fullest support, encouragement and guidance throughout this project. I was truly privileged to have the opportunity work with him. For technical assistance and additional guidance throughout this project, I would also like to thank Dr. Dauenpen Meesapyodsuk. Her extensive knowledge and experience were invaluable. I appreciate their trust in my capability to conduct the research.

I would also like to extend my appreciation to my committee members, Dr. Tak Tanaka and Dr. Nicholas Low from Department of Food and Bio Product Sciences and Dr. Aaron Beattie and external examiner Dr. Kirstin Bett from Department of Plant Science. I am thankful to all the advice and guidance offered to ensure the success of the project

I gratefully acknowledge the financial support of the Department of Food and Bio Product Sciences, University of Saskatchewan and Saskatchewan Agriculture Development Fund. Also I wish to acknowledge all the faculty and staff members in the Department of Food and Bio Product Sciences.

I would especially like to thank all my labmates for their kind supports and for being wonderful friends. Finally, I cannot continue without acknowledging the patience and understanding of my family and friends from my personal life. Special thank goes to my husband Dilshan Benaragama and loving daughter Thenuki Benaragama for their understanding, love and support given throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

PXG	Peroxygenase
AsPXG1	Oat first peroxygenase gene
EST	Expressed-Sequence Tag
LDL	Low-density lipoprotein
AA	Arachidonic acid
EPA	Eicosapentaenoic acid
DHA	Docosahexaenoic acid
DGDG	Digalactosyl diacylglycerol
FFA	Free fatty acids
TAG	Triacylglycerol
UFA	Unusual fatty acids
DGAT I/DGAT II	Diacylglycerol acyltransferase
PDAT	Phospholipid: diacylglycerol acyltransferase
PVC	Polyvinylchloride
EFA	Esterified fatty acids
TFA	Total fatty acids
PXG4	Fourth peroxygenase
EFA27	ABA-responsive gene in rice
ATS1	Arabidopsis embryo specific gene
cDNA	Complementary DNA
AtPXG1/AtPXG2/AtPXG4	Arabidopsis peroxygenase
RACE	Rapid amplification of cDNA ends
PCR	Polymerase chain reaction
RT-PCR	Reverse transcription PCR
AsPXG2	Oat second peroxygenase gene
AsPXG3	Oat third peroxygenase gene
AsPXG4	Oat fourth peroxygenase gene
ORF	Open reading frame
CORE	Collaborative Oat Research Enterprise

NCBI	National center for biotechnology information
UTR	Untranslated region
LD	Lipid droplets
ER	Endoplasmic reticulum
GOI	Gene of interest
CIAP	Calf Intestinal Alkaline Phosphatase
LB	Luria-Bertani medium
IPTG	Isopropyl-b-thiogalactopyranoside (IPTG)
YPD	Yeast extract peptone dextrose medium
BGM	Buffered minimal glycerol medium
BMM	buffered minimal methanol medium
FAME	Fatty acid methyl ester
GC	Gas Chromatography
H1/H2/H3	Three Histidines
M1-M12	Twelve Mutants
NRC	National Research Council of Canada
R1-R12	Reverse primers
F1-F12	Forward primers
wt	Wild type

1. INTRODUCTION

Oat (*Avena sativa* L.) is an annual cereal crop ranking sixth in world cereal production with more than 24 million tons annually (Reynolds, 2004). Compared to other cereals, oat is more versatile and can be used for human food, animal feed and non-food bioproducts due to its unique grain attributes with various nutritional components. Over the last 25 years, consumption of oat as a functional food has dramatically increased due to the presence of a number of compounds with recognized health benefits beyond basic nutrition. Oat is a very good source of soluble fiber (especially β-glucan), essential amino acids (particularly lysine, leucine, isoleucine), essential fatty acids (linoleic and linolenic acid), vitamins, minerals (phosphorous and iron), and other phytochemicals (avenanthramides) (Youngs and Forsberg, 1987; Matz, 1991; Marlett, 1993; Kapica, 2001; Zing, 2004; Kurtz and Wallo, 2007; Meydani, 2009). A high content of oil dispersed throughout the starchy endosperm is also interesting, not only for its quantity, but also for the unique fatty acid composition (Hamberg and Hamberg, 1996; Leonova et al., 2008). In addition to the healthy polyunsaturated fatty acids (Health Canada, 2012), oat seed oil also contains several unusual oxygenated fatty acids (Hamberg and Hamberg, 1996; Leonova et al., 2008) whose biosynthesis is the main interest in the present study.

Oat is known to produce several epoxy fatty acids, a group of oxygenated fatty acids, in seeds synthesized through a peroxygenase pathway (Meesapyodusk and Qiu, 2011). Although epoxy fatty acids have a number of potential industrial applications, for example, for the production of plastics, nylon, adhesives and paints (Blee et al., 2012), these oxygenated fatty acids can contribute adversely to the flavor of food from oats as they impart a bitter flavor mainly due to the increasing content of volatile oxidation products (Doehlert et al., 2010). The off-flavor attributes derived from the oxygenated fatty acids can adversely influence the healthy image of oat-based foods and eventually the demand of oat products. Thus, development of an effective strategy that can minimize the formation and/or accumulation of epoxy fatty acids while retaining its healthy components in oat would be important for oat breeding and oat food industry in the future.

The first peroxygenase (PXG) gene, *AsPXG1*, involved in the biosynthesis of epoxy fatty acids through a peroxygenase pathway was recently identified and characterized from oat seeds (Meesapyodusk and Qiu, 2011). However, bioinformatic analysis of oat transcript sequences indicates there might be several other PXG genes in the oat genome involved in the biosynthesis of epoxy fatty acids with different enzymatic properties. Thus, this thesis aims to molecularly clone and functionally characterize the remaining PXG genes involved in the biosynthesis of epoxy fatty acids in oats. The study consists of three specific objectives, 1) molecular cloning and sequence analysis of fatty acid peroxygenase genes in the oat genome, 2) functional analysis of putative PXG genes, and 3) mutagenesis of PXG genes for structure-function analysis.

The research started with searching oat Expressed-Sequence Tag (EST) databases prepared from developing seeds, leaves, roots and germinating seeds for homologous genes using the previously identified AsPXG1 as a query sequence, and cloning these putative *PXG* genes from oat into an intermediate plasmid. After sequence confirmation, the candidate genes were sub-cloned into a bacterial/yeast expression vector for functional analysis. The lysates of transformants were used as enzyme sources for *in vitro* assays and the function of the genes was established by analyzing fatty acid products from the assays by gas chromatography (Doehlert et al., 2010; Meesapyodusk and Qiu, 2011). The substrate specificity of an oat PXG was determined by using a variety of free unsaturated fatty acids as substrates in *in vitro* assays. Site-directed mutagenesis is one of the most important techniques used to introduce specific mutations into a gene. This technique was used in this research to investigate the structure-function relationship of an oat PXG gene. The single point mutations were introduced through an overlapping PCR approach and *in vitro* enzymatic assays of these mutants were used to identify crucial residues of the PXG enzyme for its enzymatic properties.

The information obtained from this research contributes to not only our further understanding of the biosynthesis of epoxy fatty acids through the peroxygenase pathway in oat, but also provides new knowledge on the catalytic information of this hemoprotien enzyme. The results could aid development of functional molecular markers for oat breeding to manipulate the PXG activity and epoxy fatty acid content.

2. LITERATURE SURVEY

2.1 Oat as an important cereal crop

Oat is a well-known cereal crop and ranks sixth position in world cereal production following wheat, maize, rice, barley and sorghum. Oat is well adapted to a wide range of soil types such as acidic (pH 4.5) and alkaline soils (pH 8.5), but with acidic soils being preferred. Oat is sensitive to hot and dry weather during the flowering and seed-filling period, thus it is mostly grown in cool moist climates. Therefore, oat production in the world is generally concentrated between latitudes 35-65°N and 20-46°S (Stevens et al., 2004).

Oat was brought from Europe to North America in 1565 and to Canada in 1617 (Coffman, 1977; Schrickel, 1986). Since then, oat has gradually become a major cereal crop in Canada, currently being cultivated on about 1.3 million hectares of land (Statistics Canada 2012). Oat represents 6% of the total production and export of cereal and oilseed crops in Canada. On the global scale Canada accounts for 45 to 50% of the world oat exports (Agriculture and Agri-Food Canada, 2010). Canada exported oats with a value of $620 million in 2012 to a number of key markets, and will continue to be a global leader in oat production, milling and trade. Throughout Canada, Saskatchewan is the largest producer of oats among major growing areas in the Western Prairie Provinces (i.e., Saskatchewan, Alberta, and Manitoba) (Statistics Canada 2009; Agriculture and Agri-Food Canada, 2012).

The *Avena* genus belongs to the Graminiae family with a single set of the chromosome number of seven. The genus consists of 31 known species and among them four are cultivated (Loskutov, 2001; Carlsson et al., 2007). This genus also contains three ploidy levels: diploid, tetraploid and hexaploid. The *A. sativa* cultivars are considered as the common oats and by far, are the most widespread. It is believed that cultivated oats and the wild oat species of *A. fatua* probably evolved from *A. sterilis* (also a hexaploid species) (Coffman, 1977; Li et al., 2000). However, Rajhathy and Thomas believe that *A. sativa* evolved from wild *A. fatua* and *A. sterilis* (Rajhathy and Thomas, 1974). Thus, important traits such as higher protein and oil content, large grain size, disease and cold resistance in *A. sativa* are probably derived from *A. fatua* and *A. sterilis* (Loskutov, 2001).

Despite a wide range of non-food applications such as cosmetics, pharmaceuticals and neutraceuticals, oats have traditionally been grown for two main purposes, feed and food. Although world oat production has been stable for the last five to ten years, the demand for oat as a food has gradually risen because of the recognized healthy value of the oat grain (Agriculture and Agri-Food Canada, 2010). Oat has been called 'nature's functional food' as oat kernels and oat products impart a number of health benefits beyond basic nutrition (Kapicac, 2001). Oat contains a high amount of total dietary fibers (15-22%) and comprises more soluble fibers such as beta-glucans (6.6-10.4%) than any other grain (Marlett, 1993).

Table 1: Types of minerals, vitamins and amino acids in oats

Oats: Nutritional value per 100 g

Minerals		Vitamins		Amino acids	
Calcium	54 mg	Vitamin C	0 mg	Tryptophan	0.234 g
Iron	4.72 mg	Thiamin (B1)	0.763 mg	Threonine	0.575 g
Magnesium	177 mg	Riboflavin (B2)	0.139 mg	Isoleucine	0.694 g
Phosphorous	523 mg	Niacin	0.961 mg	Leucine	1.284 g
Potassium	429 mg	Pantothenic acid	1.349 mg	Lysine	0.701 g
Sodium	2 mg	Vitamin B-6	0.119 mg	Methionine	0.312 g
Zinc	3.97 mg	Total folate	56 mcg	Cystine	0.408 g
Copper	0.626 mg	Vitamin B-12	0 mcg	Phenylalanine	0.985 g
Manganese	4.916 mg	Vitamin A	0 IU	Tyrosine	0.573 g
		Retinol	0 mcg	Valine	0.937 g
				Arginine	1.192 g
				Histidine	0.405 g
				Alanine	0.881 g
				Aspartic acid	1.448 g
				Glutamic acid	3.712 g
				Glycine	0.841 g
				Proline	0.934 g
				Serine	0.750 g

Data source: USDA National Nutrient Database
(United States Department of Agriculture)

Beta-glucan is a predominant component of endosperm cell walls in oats (Figure 1) (Buckeridge et al., 2004; Brennan and Cleary, 2005) and its consumption has been shown to reduce the risk of cardiovascular diseases by lowering the LDL cholesterol (harmful) that correlates with the build-up of plaques in blood vessel walls and blood vessel blocks. As

suggested previously, the cholesterol-lowering property of β-glucan is attributable at least in part to inhibition of cholesterol absorption from the gut (Delany et al., 2003; Errkila et al., 2005). Beta-glucan increases viscosity in the gastrointestinal tract by forming a viscous layer in the small intestine, which inhibits intestinal uptake of dietary cholesterols and re-absorption of bile acids. This inhibitory action of oat β-glucan, in turn, can increase the endogenous synthesis of bile acids from cholesterols and reduce the circulated LDL cholesterol level in the body (Othman et al., 2011).

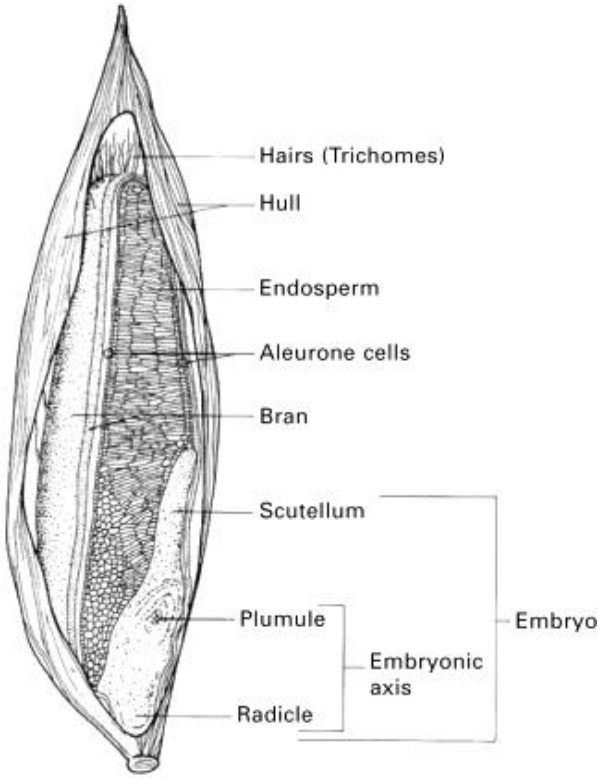


Figure 1. Anatomic illustration of an oat kernel (longitudinal section). Oat storage proteins can be found in three tissues: the embryo, the aleurone layer in the bran, and the endosperm. Mixed linkage β-glucan is a component of endosperm cell walls. Endosperm accounts for the most of the lipid accumulation in oats. Adapted from Youngs, 1986 and Zwer, 2010.

In addition to dietary fiber, oat also contains many other health promoting components such as minerals and vitamins. It is also a very good source of essential amino acids (Table 1).The mineral content of oat is generally 2-3% with mainly P and K, and lesser amounts of Mg and Ca. Compared to the other cereals, oat contains a high level of thiamin and pantothenic acid which are considered as essential nutrients (Table 1). Vitamin E, riboflavin and folic acid

important for cell signaling, proper development and function of the skin and blood cells, and rapid cell division and growth are also detected in high levels in oats (Youngs and Forsberg, 1987; Zing, 2004).

Oat also contains phytochemicals that possess a phenolic moiety with free-radical scavenging capability. These phenolic compounds exhibit anti-oxidant properties as well as anti-atherogenic and anti-inflammatory properties that can provide protection against degenerative diseases such as cancer and heart diseases (Dyke and Rooney, 2007). Among these phenolics, avenanthramides (oat flakes: 26–27 μg/g, oat bran: 13 μg/g) are exclusively present in oats and have shown positive effects in blood vessel dilation and blood pressure reduction in humans. In addition, they have also proven to have anti-itching and anti-irritation activity and are used as a remedy for the treatment of poison ivy, sunburn, eczema, and psoriasis (Kurtz and Wallo, 2007; Meydani, 2009).

Oat protein has higher nutritive value than all other cereals. The amino acid balance of oat protein is superior to those of most other grains, as its primary storage protein is globulin. Oat is the only cereal containing a globulins, a legume like protein as the major storage protein with 52% to 75% of the total proteins in groats (Zwer, 2010) (Figure 1). Globulins have a higher concentration of lysine and other essential amino acids than do typical cereal storage proteins, such as prolamins (Deshpande and Damodaran, 1990). Compared to other cereals such as wheat and barley, oats may contain twice the amount of lysine representing about 3.2–5.2% of the total protein in oat groats (Fulcher, 1986; Lockhart and Hurt, 1986).

Interestingly, oats are known to have a high content of oil with two essential fatty acids among cereals (Lehtinen and Kovirta, 2011). The total lipid content in oats varies typically between 6% and 18% of the whole grain (Butt et al., 2008; Heneen et al., 2008). According to fluorescence, scanning and transmission electron microscopies, most of the lipids are stored in the sub-aleurone and endosperm cells of the grain (Bannas et al., 2007; Heneen et al., 2008: Heneen et al., 2009). Therefore, in addition to the soluble fibers and phytochemicals, a high amount of essential fatty acids-containing oil is also nutritionally attractive for oats. Therefore, regular consumption of oat foods is recommended to reduce the risk of certain diseases and ensure a healthy life in humans.

2.2 Oat lipid and fatty acid composition

Cereal grains have low oil content in general, and the oil is mostly confined to the embryo and scutellum. However, oat is unusual in that it contains a high amount of oil in the grain. Oat can accumulate up to 18% lipids, mostly in the subaleurone cells and in the endosperm cells in the vicinity of the embryo. At the early developmental stage of an oat grain, oil is not evenly distributed and most frequently localized in the sub-aleurone cells. At the later stage of seed development, more oil accumulates in the endosperm cells near the embryonic axis (Figure 1) (Banas et al., 2007; Heneen et al., 2009). The energy value of fat is more than twice that of carbohydrate and protein. Therefore, compared with other cereals such as wheat, corn, brown rice, rye and barley, high oil content in oat can contribute to a high energy value in oatmeal.

The fatty acid composition of oat lipids is significant for the nutritional and chemical properties. According to previous studies, oat lipids consist of two types of lipid classes, neutral lipids such as triacylglycerols, diacylglycerols, free fatty acids and sterol esters, and polar lipids such as phospholipids and glycolipids (Lehtinen, 2011). Oat is rich in polar lipids where phospholipids range from 5-26% and glycolipids account for 7-12% of the total lipids (Sugawara and Miyazawa, 1999). The total lipids consist of a combination of saturated fatty acids (C14:0, C16:0, C18:0), monounsaturated fatty acids (C18:1, C20:1) and polyunsaturated fatty acids (C18:2, C18:3). Among these fatty acids, three major fatty acids that account for 90-95% of total fatty acids in oats are palmitic (C16:0), oleic (C18:1), and linoleic acid (C18:2). Others such as stearic (C18:0), linolenic (C18:3, 2-5%) and myristic (C14:0), and eicosenoic acids (C20:1, 0.5-2%) are minor fatty acids (Lehtinen, 2011; Lehtinen and Kovirta, 2011). Of the major fatty acids, only oleic acid is positively correlated with the oil content, thus an increase in the oil content also leads to an increase in the level of oleic acid in oat (Lenova, 2007).

Composition of phospholipids in oat is not unusual as compared to those in other cereals. Phosphatidylcholine is the main component of oat phospholipid (>50%) followed by phosphatidic acid (18%), phosphatidylinositol (10%) and phosphatidylethanolamine. However, composition of glycolipids in oat is very different from those of other cereals. In addition to di, tri, and tetragalactosyl diacylglycerols, oat kernels also possess mono-, di-, and tri-acyl estolides in galactolipids (Doehlert et al., 2010), that are unique to oat.

Fatty acid composition of oil is a vital determinant of oil quality for human consumption. It has been noted that a diet with low amounts of saturated fatty acids, *trans* fatty acids and

cholesterol, and a high amount of unsaturated fatty acids is more desirable for human consumption (Lichtenstein et al., 1998; Holland et al., 2001). Accordingly, composition of oat oil is considered relatively healthful due to the presence of a comparatively high level of mono-unsaturated fatty acids such as oleic acid (Peterson and Wood, 1997). Oat oil is higher in oleic acid than commonly used sunflower and soybean oils. Oat is also a good source of two essential fatty acids for humans: linoleic and α-linolenic acids that can act as precursors in the biosynthesis of very long chain ω-6 fatty acids such as arachidonic (AA) and very long chain ω-3 fatty acids such as eicosapentaenoic (EPA), and docosahexaenoic (DHA) acids, respectively. These very long chain polyunsaturated fatty acids are important for prevention of pathologies such as heart disease, arthritis, and inflammatory and autoimmune diseases in humans (Ruxton et al., 2007). However, currently oat oil is not extracted as an edible oil, mainly due to economic reasons and low oil content compared to oilseed crops (Zwer, 2010), though it may be feasible through development of high oil cultivars with good agronomic traits in the future.

Additionally, oat polar lipids, glycolipid fractions in particular, have functional properties useful in pharmaceuticals, cosmetics and food applications. Countries such as Sweden, Canada and USA are engaged in research and development of products based on oat polar lipids. In the pharmaceutical industry, liposomes from oat digalactosyl diacylglycerol (DGDG) are used as carriers to deliver pharmaceuticals. Further, DGDG is used to produce an oil-in-water emulsion cream that has extended moisturizing effects. In the food industry, oat glycolipids together with vegetable oils are used to make emulsions that can increase satiety. These kinds of products are useful in weight management programs (Peterson, 2004).

Oat lipids are also implicated in the flavor attributes of oat foods. The contribution of lipids and their roles as flavor generators have been extensively investigated in the past (Heydanek and McGorrin, 1981; Zhou et al., 1999). Oat groats when collected from the field lack flavor and their products retain a flat, green taste. Lipid oxidation products and *N*-heterocyclic compounds formed during heat processing of the groats confer the distinct flavor of the oat products (Heydanek and McGorrin, 1986).

Apart from their health benefits, unsaturated fatty acids such as oleic, linoleic, linolenic and eicosenoic acid together contribute up to 80% of the total fatty acids in oats that can undergo oxidation and rancidity reactions. Particularly polyunsaturated fatty acids such as linoleic and linolenic acids are highly susceptible to peroxidation, a process generally associated with

negative food flavors (Holland et al., 2001). However, this risk can be attenuated by the presence of high palmitic acid, a major saturated fatty acid in oats, thereby increasing oil stability.

The hydrolytic process by which free fatty acids (FFAs) are generated from triacylglycerols is believed to be initiated by endogenous enzymes, lipases. As lipases are liberated by damaged kernels, oat lipids in damaged or improperly handled oats can undergo dramatic changes during storage, such as, extensive hydrolysis due to the increased surface area and contact between the enzyme and the lipid substrate. Therefore, prevention of lipid hydrolysis should be a primary goal in the manufacture of non-deteriorated oat products (Liukkonen et al., 1992). According to Welch (1977), moist bruised grains yield FFA levels up to 16% of oil after seven months of storage, in comparison with 4% in sound dry grains. Moreover, levels of FFAs at 30–40% are reported in experimentally processed oats even when lipase has been inactivated (Ekstrand et al., 1992). Furthermore, high temperature and high humidity during storage can accelerate FFA accumulation in ungrounded oats. For example, an increase of FFAs is observed in unground oats during storage at 30 ºC and 80% relative humidity (Molteberg et al., 1996). Interestingly, it is found that lipids in intact groats (un-processed/un-damaged), stored at 20 ºC and 12–14% moisture, remain stable. Therefore, to prevent the release of free fatty acids from glycerolipids for oxidation, oat food processing often includes a lipase-inactivation step as a precaution, with brief steam treatment at 90–100 °C (Zhou et al., 1999).

Oat triacylglycerol (TAG) contains five dominating fatty acids: palmitic, steraric, oleic, linoleic and linolenic acids, as well as several unusual fatty acids with different chain length (i.e., greater than 18 carbons or shorter than 14 carbons) and position or number of double bonds or presence of functional groups such as epoxy and hydroxyl groups (Napier, 2007). The epoxy fatty acids have also been found in other plant species, for instance, a high level of vernolic acid, *cis*-12,13-epoxy-octadec-*cis*-9-enoic acid, is found from the seed oil of *Vernonia anthelmintica.* In addition, *Alchornea cordifolia*, *Tragopogon porrifolius* and *Chrysanthemum coronarium* accumulate 14-epoxy, cis-11-eicosenic, 9-epoxy-octaecanoic and 9-epoxy, cis-12-octadecenoic acids, respectively, in their seeds (Li et al., 2010).

2.3 Epoxy fatty acids in plants

Although plant cell membranes are primarily composed of five common fatty acids, i.e. stearic, palmitic, oleic, linoleic and linolenic acids, storage lipid triacylglycerols (TAGs) in seeds can have a wider range of fatty acids including unusual fatty acids. Unlike most commercial oilseed crops, many exotic plant species can accumulate high levels of unusual fatty acids (UFA) such as hydroxy, epoxy and acetylenic fatty acids in seed TAGs (Rezzonico et al., 2004). For example, *Vernonia galamensis, Euphorbia lagascae* and *Stokesia laevis* are prominent in that they can accumulate 60–80% of an epoxy fatty acid known as vernolic acid (cis-12-epoxyoctadeca-cis-9-enoic acid) in seeds (Bafor et al., 1993; Li et al., 2010).

The discovery of unusual fatty acids in seed oils has attracted considerable attention regarding their biosynthetic mechanisms and industrial applications. Accumulation of unusual fatty acids in TAGs requires at least two major catalytic steps: biosynthesis of unusual fatty acids by insertion of a functional group into an acyl chain and incorporation of unusual fatty acids into TAGs (Rezzonico et al., 2004). For instance, biosynthesis of vernolic acid in *E. lagascae* and *Vernonia galamensis* results from epoxidation of the Δ12 double bond of linoleic acid by an epoxygenase (Bafor et al., 1993; Liu et al., 1998; Cahoon et al., 2002). Epoxidation of fatty acids in plants can be catalyzed by three different mechanisms (Meesapyodsuk and Qiu, 2011). The first one identified by Stymne and colleagues (Lee et al., 1998) involves the synthesis of vernolic acid: 12,13-epoxy-18:1-9c in *Crepis palaestina* catalyzed by a Δ12 desaturase-like epoxygenase with characteristics similar to the membrane desaturase containing non-heme iron. The second mechanism involves biosynthesis of vernolic acid in *Euphobia lagascae* by cytochrome p450-like epoxygenase (Lee et al., 1998; Cahoon et al., 2002). The peroxygenase pathway is the third mechanism underlying the biosynthesis of epoxy fatty acids in oats (Meesapyodsuk and Qiu, 2011) where a peroxygenase catalyzes hydroperoxide-dependent epoxidation of unsaturated fatty acids (Hamberg and Hamberg, 1990; Blée et al., 1993). Oat seeds have a high capacity to synthesize epoxy-hydroxy/epoxy acids from unsaturated fatty acids such as oleic, linoleic and a-linolenic acids through a peroxygenase pathway (Hamberg and Hamberg, 1996). Some plants produce epoxy fatty acids in response to pathogen attack (Blee, 1998), however, they do not accumulate epoxy-triglycerides in their seeds. For accumulation of epoxy fatty acids in TAGs, other enzymes besides epoxygenase, such as diacylglycerol acyltransferase (DGAT I and DGAT II), and phospholipid: diacylglycerol acyltransferase

(PDAT) involved in the assembling process are required. Accordingly, these enzymes have important roles in epoxy fatty acid production of *Vernonia galamensis*, *Euphorbia lagascae* and *Stokesia laevis* (Li et al., 2010).

To our knowledge, none of the species mentioned above that produce epoxy fatty acids is cultivated on a large scale for industrial uses. Vegetable oils such as linseed and soybean oil are currently utilized to produce epoxidized oil by introducing epoxy group across the double bonds of polyunsaturated fatty acids. However, the chemical modification of vegetable oils is expensive and requires the use of a catalyst (Perdue et al., 1986). Because of potential industrial applications of epoxy fatty acids, there has been substantial research over the last decade attempting to obtain commercially viable crops with high levels of epoxy fatty acids especially vernolic acid in the seed oil (Cahoon, 2002; Napier, 2007; Li et al., 2010). Most of the research focuses on transferring epoxy fatty acid biosynthetic machinery from the wild plants to a high yield oilseed crop. For the genetic engineering, a number of genes encoding enzymes responsible for biosynthesis of epoxy fatty acids in plants are cloned and characterized (Thomeaus et al., 2001). Expression of P450-like epoxygenase from *Euphorbia lagasceae* under a seed-specific promoter in soybean somatic embryos results in nearly 8% (w/w) vernolic acid (Cahoon et al., 2002) in the embryos, while expression of Δ-12 epoxygenase gene from *Crepis palaestina Cpal2* under a seed-specific napin promoter produces 6.2% epoxy fatty acids in *Arabidopsis thaliana* seeds (Singh et al., 2001). Furthermore, co-expression of diacylglycerol acyltransferase (VgDGAT2) from *Vernonia galamensis* and an epoxygenase from *Stokesia laevis* (*SIEPX)* in soybean leads to accumulation of 25% of the industrial-valued vernolic acid in mature seeds (Li et al., 2010).

Epoxy fatty acids have properties suitable for industry use (Budziszewski et al., 1996), as the oxirane moiety of an epoxy molecule is highly reactive. Epoxy fatty acids are used as a plasticizer of polyvinylchloride (PVC), and as a precursor in the manufacture of nylon. The ability of epoxy groups to crosslink with each other makes epoxy fatty acids-containing oils useful in industries that produce adhesives and paints. Furthermore, epoxidized oils can be used in the production of coatings with low levels of volatile organic compounds (Budziszewski et al., 1996). In plants, epoxy fatty acids are not only confined to the seeds, they are also components of cutin, the scaffold of cuticles that cover all the aerial surface parts of plants. The plant cutin is composed of networks of inter-esterified hydroxyl and epoxy derivatives of C16 and C18 fatty

acids, in which epoxy fatty acids function as monomers in the polymers (Budziszewski et al., 1996; Lequeu et al., 2003). More interestingly, epoxy fatty acids are also recognized to act as precursors for the synthesis of oxylipins in response to biotic and abiotic stress in plants. Plant oxylipins are a group of crucial signaling molecules for coordinated metabolisms in response to developmental information and environmental stimuli (Chehab et al., 2007).

2.4 Epoxy fatty acids in oat

Oat is known to produce several epoxy fatty acids in seeds that account for 0.5-5% of the total fatty acids. These include 9,10-epoxy-18:0, 9,10-epoxy-18:1-12*c* and 12,13-epoxy-18:1-9*c* and they are mainly present in neutral lipids such as triacylglycerols and diacylglycerols (Figure 2). Although these compounds have recently been independently identified in oats by gas chromatography-mass spectrometry, accurate quantity of individual epoxy fatty acids in different lipid classes is yet to be determined (Leonova et al., 2008: Doehlert et al., 2010; Meesapyodsuk and Qiu, 2011).

Fatty acid metabolism in plants is complex, as the biosynthetic pathway is compartmentalized in different organelles and extensive movement of the intermediates among different pools of lipid classes from organelles occurs in the cells. Ultimately, modification of fatty acids would involve many enzymes in different compartments, such as fatty acid synthase, fatty acyl thioesterase, desaturase, hydroxylase, lipoxygenase and peroxygenase (Napier, 2007; Piazza et al., 2001, 2003). Peroxygenase (PXG) is a new type of oxygenase for producing epoxidized fatty acids. In oats, several epoxy fatty acids are synthesized through a peroxygenase pathway. The pathway includes two catalytic reactions catalyzed by PXG and lipoxygenase, respectively.

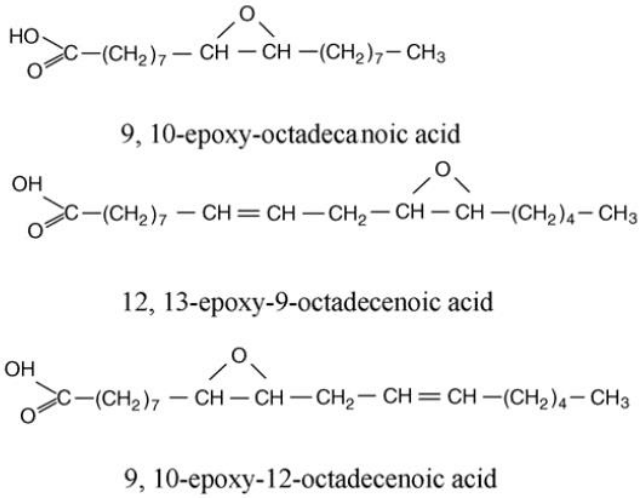


Figure 2. Structures of epoxy fatty acids identified in oat lipids. (Adapted from Doehlert et al., 2010)

The peroxygenase pathway constitutes one branch of the larger lipoxygenase pathway in oat. As the first step in the pathway, fatty acid hydroperoxides are produced by oat lipoxygenase and then epoxidation of unsaturated fatty acids (C18:1, C18:2, C18:3 or C16:3) occurs where fatty acid hydroperoxides are used as oxygen donors by a peroxygenase to oxidize unsaturated fatty acids into corresponding epoxy fatty acids (Hanano et al., 2006: Messapyodusk and Qiu, 2011). In oats, epoxy fatty acids 9,10-epoxy-18:1-12c and 12,13-epoxy-18:1-9*c* are derived from intermolecular peroxygenation, i.e. epoxidation of a double bond of a unsaturated fatty acid in presence of a fatty acid peroxide from another fatty acid (Meesapyodusk and Qiu, 2011) (Figure 3). Even though epoxy fatty acids in oat have potential for a number of industrial applications, these oxygenated fatty acids have detrimental effects on the quality/flavor and the shelf life of the food products made from oats. They are bitter in flavor and cause rancidity in oat foods. In addition, a high level of fatty acid oxidation can be implicated in some diseases such as atherosclerosis, liver damage and intestinal tumors (Lehtinen et al., 2003; Marmesat et al., 2008).

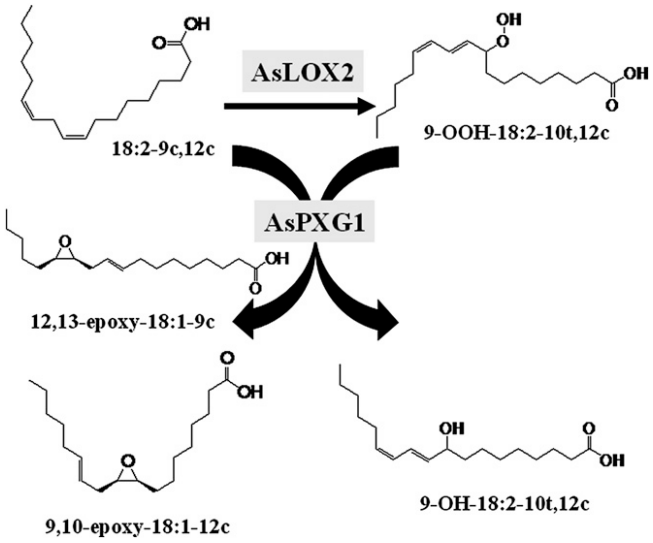


Figure 3. The peroxygenase pathway in oats. Two major epoxy fatty acids: 9,10-epoxy-18:1-12*c*, 12,13-epoxy-18:1-9*c* and a hydroxyl fatty acid: 9-OH-18:2-10*t*,12*c* as a co-product are produced when linoleic acid is incubated with oat lipoxygenase (AsLOX2) and peroxygenase (AsPXG1), respectively in an *in vitro* study. Adapted from Meesapyodusk and Qiu, 2011.

In oat processing industry, enzyme inactivation via steaming or dry heat treatment is routine in order to overcome the detrimental effect of enzymatic activity in lipid oxidation. However, oxygenated fatty acids such as epoxy and hydroxy fatty acids are found in treated oat seeds. According to Doehlert et al., 2010, epoxy fatty acids are identified in the treated oat groats and flour over a storage period of more than 22 weeks. Under extended conditions, epoxy fatty acids accumulate more in oat flour than in oat groats. These fatty acids can be found in both esterified and free fatty acid forms. Epoxy fatty acid levels are lower in the steamed, but higher in the roasted flour and groats, while higher amounts of epoxy fatty acids are found in esterified fatty acids (EFA) than in free fatty acids (FFA) in oat groats. Oat flour without the treatment contains more epoxy fatty acids than that treated in the FFA pool, but no difference is found in the amount of epoxy fatty acids in the EFA form between the treat and non-treated flour (Doehlert et al., 2010) (Table 2).

Table 2: Epoxy fatty acids in raw, steamed, and roasted oat groats and flour as percentage of free fatty acids (FFA), esterified fatty acids (EFA), and total fatty acids (FFA+EFA=TFA) (Adapted from Doehlert et al., 2010)

Treatment	Extracted from stored groats			Extracted from stored flour		
	FFA	EFA	TFA	FFA	EFA	TFA
Raw	0.47a	1.52a	1.89a	3.22a	1.77a	5.00a
Roasted	0.36b	0.90b	1.25b	3.05a	2.13a	5.18a
Steamed	0.05c	0.56b	0.61c	0.30b	1.72a	2.02a

Values in the same column followed by the same letter do not differ significantly ($P < 0.05$).

According to these findings, hydrothermal treatments done prior to the oat storage might not be completely effective in inactivation of enzymes such as oat lipoxygenase and peroxygenase involved in the biosynthesis of oxygenated fatty acids. On the other hand, enzyme inactivation via heating may promote formation of unwanted products through non-enzymatic oxidation. Thus, a more reliable and effective strategy should be implemented to ensure elimination of enzyme-based oxygenated fatty acids, thereby extending the shelf life of oat-based food products. Better understanding of the enzymes involved in the biosynthesis of oxygenated fatty acids i.e. epoxy fatty acids in particular, is therefore crucial to aid the development of a strategy for the oat industry.

2.5 Peroxygenase in biosynthesis of epoxy fatty acids in oat

Peroxygenase (PXG), a sub-class of oxidoreductases widespread in the plant kingdom. It belongs to an enzyme group EC 1.11.2.3 of the oxidoreductases (Fuchs and Schwab, 2013). It is a hydroperoxide-dependent monooxygenase defined in 1977, almost 40 years ago, by a labeling study using pea (*Pisum sativum)* microsomes. The enzyme catalyzes the transfer of an oxygen atom from a hydroperoxide to another substrate that is oxidized, which is referred to as intermolecular peroxidation. Accordingly, plant peroxygenase can catalyzes hydroxylation reactions of aromatics, sulfoxidations of xenobiotics, and most importantly epoxidations of unsaturated fatty acids (Hanano et al., 2006). In addition to that, peroxygenase can also catalyze the transfer of an oxygen atom from the hydroperoxy group of a molecule to epoxidize a double bond in the same molecule, which is called as the intra-molecular peroxidation, resulting in

formation of epoxy-hydroxy fatty acid (Figure 4) (Hanano et al., 2006; Meesapyodusk and Qiu, 2011).

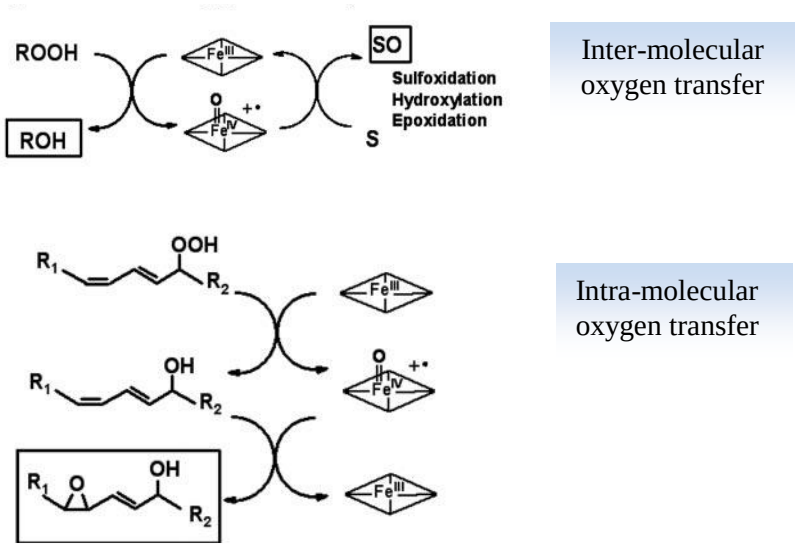


Figure 4. Inter-molecular and intra-molecular oxygen transfer reactions catalyzed by plant peroxygenase in the presence of hydroperoxides. Upper panel represents peroxygenase mediated heterolytic cleavage of the hydroperoxide, corresponding alcohol and a ferryl-oxo complex formation followed by sulfoxidation, hydroxylation and epoxidation through intermolecular oxygen transfer. Whereas, lower panel represent peroxygenase catalyzed intra-molecular oxygen transfer. Adapted from Hanano et al., 2003.

Hydroperoxide-dependent epoxidation is a new mechanism found in the biosynthesis of epoxy fatty acids and epoxy-hydroxy fatty acids in plants (Ishimaru and Yamazaki, 1977; Hamberg and Hamberg, 1990; Blee et al., 1993; Hamberg and Hamberg, 1996; Matsunaga and Shiro, 2004). In mammals, most of these reactions are performed by cytochrome P450 monoxygenases. However, plant PXG does not share any significant sequence similarity with cytochrome P450 monoxygenase. In addition, its activity differs from members of the cytochrome P450 family in the use of an oxygen donor and cofactor, as PXG does not depend on the cofactor such as NAD(P)H and molecular oxygen for substrate hydroxylation, sulfoxidation and epoxidation (Hanano et al., 2006; Fuchs and Schwab, 2013).

Structurally plant PXG belongs to a caleosin family, a group of small oil body-associated membrane proteins characterized with both calcium binding and heme-binding motifs. Therefore, PXG activity is strictly dependent on binding calcium to a helix E and F (EF)-hand

motif and ferric-heme to histidine residues. However, phosphorylation of the protein has no impact on the catalysis. In terms of enzyme activity, calcium ion is crucial in preserving a catalytically active conformation and maintaining the heme environment of peroxygenase (Hanano et al., 2006). Initially, association of the peroxygenase activity with caleosin was surprising as caleosin was first described as an oil body protein that plays a structural role in maintaining the integrity of oil bodies. However, with the advent of recent genomic sequencing, they are widely identified in fungi and other plant species localized not only in oil bodies, but also in non-lipid storage organelles and tissues (Murphy, 2005; Partridge and Murphy, 2009; Meesapyodusk and Qiu, 2011).

During the peroxygenation, a heme group bound to the enzyme catalyzes the heterolytic cleavage of a hydroperoxide O-O bond to form a ferryl-oxo complex, while the hydroperoxide is reduced to an alcohol, the oxygen atom of the ferryl-oxo complex is transferred to C=C double bonds of unsaturated fatty acids to produce corresponding epoxy fatty acids as shown in Figure 4. The enzyme is believed to be associated with oil bodies or lipid droplets via a hydrophobic central domain embedded into the monolayer of the droplet membrane, while the C and N-terminal protein regions are located on the surface (Blee et al., 1993; Hanano et al., 2006; Fuchs and Schwad, 2013).

According to their protein structure, caleosins are divided into two classes. Class I caleosins exhibit only marginal epoxygenase activity when they are heterogeneously expressed under standard conditions. They can hardly account for an efficient fatty acid epoxidation of unsaturated fatty acids in green plants (Blee et al., 2012). Examples of this class of caleosin/peroxygense are found in Arabidopsis (PXG1) and oat (PXG1) and rice (EFA27) and are mostly present in seeds and seedlings, leading to the conclusion that these enzymes are seed specific (Naested et al., 2000; Meesapyodusk and Qiu, 2011; Blee et al., 2012). In sharp contrast, Class II caleosins are not confined to seeds and they are spread throughout the plant. For instance, recently identified PXG4 from Arabidopsis is found in roots, stems, leaves and flowers of Arabidopsis indicating that peroxygenase can function in organs other than seeds. Also, unlike Class I peroxygenase, Class II peroxygenases show higher catalytic activity (Blee et al., 2012).

Hamberg and Hamberg (1996) employed oat as a model system to study enzymatic properties of PXGs. PXG-catalyzed fatty acid epoxidation in oats was first confirmed using a particle fraction of oat seed homogenate. The study showed a pronounced peroxygenase activity

in oats as measured by hydrogen peroxide-dependent epoxidation of oleic acid. Linoleic acid was converted to epoxy, alcohol epoxy, and hydroxy fatty acids in a subcellular particle fraction (Hamberg and Hamberg, 1996). In addition, not only fatty acids but also acyclic alkenes with internal double bonds and an alkene with an aromatic substituent were substrates of oat peroxygenase. Within a group of monounsaturated fatty acids or esters and *cis* alkenes tested, subtle difference in activity could be observed, as PXG activity was much reliant on the position of the double bond and the structure of the substrate (Piazza et al., 2003).

The first plant PXG gene was cloned from Arabidopsis (*Arabidopsis thaliana*) in 2006 using the sequence information of a partially purified PXG protein from oat. Surprisingly the encoded protein shares high sequence identity with caleosins (Hanano et al., 2006). Although oat has long been known to produce epoxy fatty acids in seeds by a PXG pathway, the gene encoding oat PXG was not cloned until recently (Meesapyodsuk and Qiu, 2011). The study used Arabidopsis ATS1, an embryo specific caleosin, as a query sequence to search an oat EST database to identify a cDNA *AsPXG1* encoding fatty acid PXG.

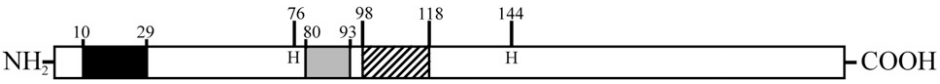


Figure 5. Primary structure of putative peroxygenase AsPXG1 from oat (*Avena sativa).* According to Meesapyodsuk and Qiu, 2011

Oat AsPXG1 catalyzes hydroperoxide-dependent monooxygenation of unsaturated fatty acids. It is a small protein with 249 amino acids belonging to Class I caleosins with conserved His residues (His-76 and His-144) and a EF-hand motif (between position 80 and 93) that are believed to be involved in heme and calcium binding, respectively (Figure 5). Similar to Class I Caleosin, AsPXG1 has a single 20-amino acid hydrophobic membrane-associated domain (between position 98 and 118) in the middle of the protein sequence for imbedding to the monolayer membrane of oil droplets. AsPXG1 shares high amino acid sequence identity to an abscisic acid-induced protein (EFA27) from rice (Oryza sativa, 66%), AtPXG1 (At4g26740, 60%) and AtPXG2 (At5g55240, 58%) from Arabidopsis (Hanano et al., 2006; Meesapyodsuk and Qiu, 2011).

The tissue specific expression study shows that *AsPXG1* is solely expressed in developing seeds, not in other tissues such as roots, leaves, glumes and germinating seeds of oat. Analysis of enzymatic properties of *AsPXG1* found that the enzyme can only use unsaturated fatty acids with double bonds in the *cis*- configuration as the substrate and oleic acid is the most preferred substrate. Unsaturated fatty acids with *trans*-configuration are not accepted as substrates by the enzyme (Meesapyodsuk and Qiu, 2011). When expressed heterogeneously, AsPXG1 catalyzes strictly hydroperoxide-dependent epoxidation of unsaturated fatty acids giving rise to corresponding epoxy fatty acids. The enzyme possesses somewhat selectivity for both substrate and co-substrate which are subjected to oxidation and reduction, respectively.

According to Partridge and Murphy (2009), there are six caleosin-like genes identified in the Arabidopsis genome (i.e., *AtCLO1*- *AtCLO6*). Preliminary bioinformatic analysis of oat transcript databases found at least six distinct caleosin-like EST sequences in oat developing seeds. These ESTs must derive from other PXG genes in the oat genome. Although oat first PXG *AsPXG1* is well characterized, enzymatic properties, substrate specificity and biological functions of other *AsPXG1*-like genes in the oat genome still remain to be determined. Molecular cloning and functional characterization of the remaining PXG genes would contribute to our further understanding of the biosynthesis of epoxy fatty acids in oat.

3. RESEARCH STUDIES

3.1 Study 1: Molecular cloning and sequence analysis of PXG genes in oat

3.1.1 Introduction

Oat seed peroxygenase exhibits high specific activity for fatty acid epoxidation. Thus, oat can serve as a model system to study enzymatic properties of peroxygenase (PXG). Identification of genes that encode peroxygenase able to catalyze the epoxidation of physiological unsaturated fatty acids is the first step toward enzymatic studies and manipulation of epoxy fatty acid levels in oat. The first gene identified in oat that encodes PXG, AsPXG1, was recently cloned and characterized *in vitro* by our laboratory (Meesapyodusk and Qiu, 2011). This study aims to identify and clone additional PXG genes involved in the biosynthesis of epoxy fatty acids in oat.

3.1.2 Abstract

The oat *AsPXG1* gene was previously confirmed to catalyze hydroperoxide-dependent monooxygenation of unsaturated fatty acids, producing corresponding epoxy fatty acids. To increase the likelihood of finding additional oat genes encoding PXG, nucleotide and amino acid sequences of this gene were used as query sequences to search oat EST databases from different tissues, which resulted in identification of many homologous sequences. Assembly of these EST sequences gave rise to six contig sequences. Two contig sequences (Contig 2 and 3) that possess high sequence similarity to the query and align well with caleosin proteins were selected for further analysis. As both contig sequences are not full length, missing 3' ends, 3' RACE (rapid amplification of cDNA ends) PCR was thus performed to obtain the missing ends of the two sequences. The full length cDNAs retrieved by RT-PCR were named *AsPXG2* and *AsPXG3*, respectively. While retrieving *AsPXG3*, an additional isoform sequence was identified from the amplification, thus named as *AsPXG4*. The open reading frame (ORF) of each putative PXG gene, i.e. *AsPXG2, AsPXG3* and *AsPXG4* was determined and subsequently cloned for the functional study. The full length ORF of *AsPXG2* is 702 bp long encoding 233 amino acids, and ORFs of *AsPXG3* and *AsPXG4* are 627 bp long encoding 208 amino acids. Sequence analysis

showed that AsPXG3 and AsPXG4 belong to Class II caleosin proteins, while AsPXG2 belongs to Class I caleosin proteins. All the AsPXGs comprise a hydrophobic domain presumably for lipid droplet anchoring, conserved hisditine residues for heme-binding and an EF-hand motif for calcium-binding.

3.1.3 Hypothesis

There are additional oat PXG sequences in oat genome that can be identified by searching oat EST databases. Sequence analysis of these sequences would reveal the structural feature and phylogenetic class of PXG/caleosin proteins.

3.1.4. Experimental approach

3.1.4.1 Homologous search for additional oat PXG sequences

Since the first peroxygenase gene in oats, *AsPXG1*, was identified in an EST database prepared from developing seeds, oat EST databases prepared from developing seeds as well as from other tissues such as leaves, roots, glumes and germinating seeds available at the Collaborative Oat Research Enterprise (CORE) and National Center for Biotechnology Information (NCBI) database were searched using *AsPXG1* as a query sequence to identify additional PXG sequences. DNAstar's SeqMan offline software (Lasergene 9 Core Suite, Madison, USA) was used for the assembly and alignment of homologous sequences and contig sequence construction. The consensus sequence of each contig was analyzed by SeqBuilder (SEQBUILDER, 2005) to define the ORF and translate it into an amino acid sequence. Amino acid sequences were aligned with the original query sequence by the MegAlign program. Primers for cloning were designed using PrimerSelect (Lasergene 9 Core Suite, Madison, USA).

3.1.4.2 Cloning of PXG genes from oat

(a) Isolation of Total RNA from oat

Total RNA from developing seeds of oat cultivar CDC Dancer was extracted using TRIzol reagent (Invitrogen Canada Inc., Burlington, Ontario). TRIzol is a ready-to-use reagent designed to obtain high quality total RNA from plant tissues within a short period of time. It can effectively maintain the integrity of the RNA due to its highly effective inhibitory action on RNase. Prior to the extraction, 200 mg of developing oat seeds was kept frozen with liquid

nitrogen and ground with a mortar and a pestle into a fine powder. To 50-100 mg of the tissue, 1mL of TRIzol reagent was added. After mixing, the sample was incubated for 5 minutes at room temperature. Then, 0.2 mL of chloroform per 1 mL of TRIzol reagent was added. The sample was capped and shaken vigorously for 15 seconds by hand. For the phase separation, the sample was incubated for 2-3 minutes at room temperature, centrifuged for 15 minutes at 4ºC at 12000xg. The aqueous phase containing the RNA was pipetted out and placed in a new tube for RNA precipitation.

For RNA precipitation, 0.5 mL of 100% isopropanol (per 1mL of TRIzol reagent) was added to the aqueous phase. After mixing, the sample was incubated at room temperature for 10 minutes, and then centrifuged at 12000xg for 10 minutes at 4ºC. After the precipitation, RNA wash was performed using 1mL of 75% ethanol (per 1mL of TRIzol reagent from the initial homogenization) with vortexing and centrifugation at 7500xg for 5 minutes at 4ºC. The ethanol was discarded and the pellet was air or vacuum dried for 5-10 minutes. The RNA pellet was re-suspended in RNase-free water for quality and quantity analysis.

The RNA concentration was measured with a DU730 Spectrophotometer by the absorbance at 260nm using the nucleic acid analysis program (Beckman Coulter Canada, Inc., Mississauga, Ontario). For RNA electrophoresis, 5 µL RNA was denatured at 65°C for 5 minutes, cooled down on ice for 2 minutes before loaded into a 1.2% (w/v) agarose gel with 2 µL ethidium bromide (10 mg/mL). The image of the RNA quality was visualized under UV light by Alpha Imager HP (Alpha Innotech Corp., Ottawa, Ontario). Remaining RNA samples were stored at -80°C.

3.1.4.3 Rapid amplification of the cDNA ends (RACE)

The RACE method was performed to obtain the missing 3' ends of the genes. The first-strand cDNA was synthesized using a SuperScript III first-strand synthesis system (Invitrogen Canada Inc., Burlington, Ontario). 5 µg of total RNA, 1 µL of oligo(dT)20 (50 uM) (Table 3), 1 µL of 10 mM dNTP, and 13 µL of water were added to a nuclease-free micro-centrifuge tube. The mixture was incubated at 65 ºC for 5 minutes and placed on ice for 1 minute. After brief centrifugation of the tube, 2 µL 10X RT Buffer, 2 µL 0.1M DTT, 4 µL of 25 Mm MgCl2, 1 µL RNaseOUT and 1 µL of SuperScriptIII RT were added. The solutions were mixed via pipetting up and down, and then incubated at 50 ºC for 50 minutes, followed by reaction inactivation by

heating to 85 ºC for 5 minutes. After the first strand synthesis, 1 µL of RNase H was added to remove the RNA template. The synthesized cDNA was then stored at -20°C and used as the template for PCR amplification.

For the 3' RACE-PCR amplification, two gene specific primers were designed according to the 3' end sequence information of Contig sequence 2 or 3. IB1 outer primer and IB2 inner primer were designed for Sequence 2, whereas IB3 outer primer and IB4 inner primer were designed for Sequence 3 (Table 4). These inner and outer nested primers were nearly 100 bp apart from each other at the 3' end. Using these gene-specific and common RACE primers (PTO27 and PTO26), RACE PCRs were performed to amplify the missing ends.

Table 3: Common primers for cDNA synthesis and 3' RACE.

Primer Name	Sequence 5' to 3'	Primer Description
PT028/ Oligo(dT) $_{20}$	GCGAGCACAGAATTAATACGACTCACTATAGGT12VN	Adaptor primer
PT027	GCGAGCACAGAATTAATACGACT	3' RACE outer primer
PT026	CGCGGATCCGAATTAATACGACTCACTATAGG	3' RACE inner primer

Table 4: Gene specific primers for 3'RACE and full length retrieval of AsPXG2 and AsPXG3.

Target gene	Primer sequence from 5' to 3'	Primer Description
Contig sequence 2 (partial)	CCACAAGGGCAAGCATGGAAGTGATTC	GS outer primer- IB1 (Tm= 66 ºC)
	CAGAAGATGAGATTAGCACCATGC	GS inner primer- IB2 (Tm= 54.7 ºC)
Contig sequence 3 (partial)	GCAAGCATGGGAGT GATAC AGACACG	GS outer primer- IB3 (Tm= 62 ºC)
	TAAAGCAAATAGGGAGCGGAATG	GS inner primer- IB4 (Tm= 56.5 ºC)
AsPXG2	GTGAATTCTTATGTCGTCGTCGTCTG	GS forward primer- IBW1 (Tm= 58 ºC)
	GCTCTAGAGCTGAATTCAGGAAGATTTC	GS reverse primer- IBW2 (Tm= 58.7 ºC)
AsPXG3	GAATTCATGGCCTCGAAGCCCGC	GS forward primer- IBW3 (Tm= 66 ºC)
	GAATTCTCATTTCTTCTTGGCAGAATCC	GS reverse primer- IBW4 (Tm= 60 ºC)

GS: gene specific. All primers were synthesized by Sigma-Aldrich Canada Ltd. Oakville, Ontario. Restriction sites, start and stop codons were underlined.

For initial PCR amplification, 0.5U of Phusion® polymerase (Finnzymes, 2012) was used in a 25µL reaction, with 200 ng of template cDNA, 5X Phusion HF buffer, dNTP and primer concentrations of 200 µM and 0.5 µM, respectively (New England Biolabs Inc., 2013). The reaction was held at 98°C for 30 seconds, followed by the first 35 cycles of denaturation at 98°C for 10 seconds, annealing at 55 °C for 30 seconds, and elongation at 72°C for 30 seconds. The reaction was then held at 72°C for 10 minutes. Larger fragments obtained from the initial PCRs agarose gel were used as the templates in the nested PCRs. The nested PCR reaction was loaded on a 1.0% (w/v) agarose gel with 2 µL ethidium bromide (10 mg/mL) and checked under UV light by Alpha Imager HP. The amplified fragments were cut and gel purified and ligated into a TA-cloning vector pGEM-T (Promega, Nepean Ontario) (Figure 6). The vector was linearized with a 3' terminal thymidine to both ends. The 3' T overhangs at the insertion site would improve ligation efficiency of a PCR product with 5' A overhangs by preventing re-circularization of the vector. For cloning each nested fragment into the pGEM-T vector, a single-nucleotide dATP-overhang was added to the purified inserts. Gel-extracted fragment was transferred to a 0.2 mL tube, and mixed with 5 µL of Invitrogen's 10X PCR buffer (-MgCl2), 1.5 µL of 50 mM MgCl2, 1 µL of 10 mM dNTPs, and 0.25 µL of Invitrogen Taq polymerase. The reaction mixture was incubated for 20 minutes at 72°C. The inserts with dATP overhangs was cloned into pGEM-T vector using a Promega's kit. 1 µL of pGEM-T linearized vector was ligated at 4°C overnight with 3 µL of insert with dATP-overhangs in a mixture containing 1 µL of T4 DNA (Invitrogen, 2002) and 5 µL of 2X rapid ligation buffer. Recombinant vectors carrying inserts were transformed into electrocompetent TOP10 *E. coli* cells using an electroporator, set at 1800 mV, and using glass cuvettes with 1mm gap.

For each transformation, 50 µL of TOP10 cells were thawed for 3 minutes on ice, and 2 µL of the overnight ligation of the recombinant vector was added to these cells, and gently mixed. The 52 µL volume was transferred into the cuvette, and set on ice for 5 minutes. After electroporation, 1 mL liquid LB medium was added into the cuvette. The total sample was then transferred into a plastic 15mL tube, incubated at 37°C for 2-3 hours, and plated on a LB agar plate containing 40 µL of X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside) and 40 µL of 0.1M of IPTG (isopropyl β-D-1-thiogalactopyranoside).

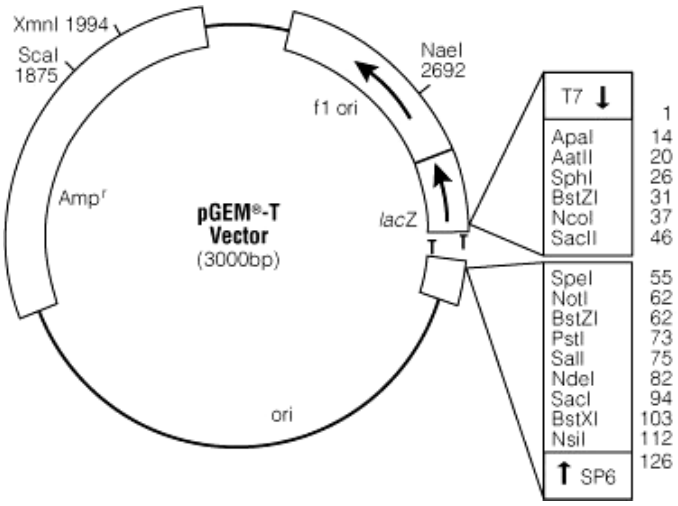


Figure 6. The pGEM-T vector circle map (Reproduced from technical manual of pGEM-T vector systems, Promega corporation, Madison, USA).

The plates were incubated at 37°C overnight, after which several white colonies were selected and screened via PCR. The PCR screening was performed to confirm the presence of a insert in the intermediate vector. For this, the same gene-specific primer along with a vector-based primer was used in a PCR reaction. Transformants carrying the inserts were selected; the plasmids were extracted, and sent for sequencing. Gene specific primers containing start and stop codon along with restriction enzyme EcoR1 sequence, were designed to obtain full length putative genes (Table 4). The gene names *AsPXG2* and *AsPXG3* were given for Sequence 2 and 3, respectively.

3.1.4.4 Sequence analysis of *AsPXGs*

Gene specific primers were used to retrieve the full-length cDNAs using reverse transcriptase-PCR (Invitrogen Canada Inc., Burlington, Ontario). Phusion DNA polymerase (0.25 µL) (Finnzymes Canada Inc., Ottawa, Ontario) was used in 25 µL PCR amplification reaction. The PCR reaction was loaded on an agarose gel, the fragments with expected size were cut and gel purified and ligated into cloning vectors pGEM-T (Promega, Nepean Ontario) as described before. The recombinants were then transformed into *Escherichia coli (*TOP10) (Invitrogen Canada Inc., Burlington, Ontario). Positive transformants were selected on LB plates

with IPTG and X-gal, the plasmids were extracted from the transformants and sent to PBI, National Research Council of Canada, Saskatoon, Saskatchewan for sequencing.

Once the full length nucleotide sequence was confirmed, amino acid sequence of the ORF was determined by translating the nucleotide sequence in DNAstar's SeqBuilder. The protein sequences were then aligned with the query sequence to search for the presence of conserved regions i.e. heme-binding residues and a calcium-binding motif, required for the activity of the oat peroxygenase. Meanwhile the sequencing results were also used to search against the NCBI non-redundant protein sequence databases for homologous peroxygenase proteins from other species.

Membrane-spanning regions were predicted using the OCTOPUS topology predictor (Stockholm University, Stockholm Bioinformatics Center, 2008). This online tool predicted the preference for each residue to be located in a transmembrane (M), interface (I), close loop (L) or globular loop (G) environment and the preference for each residue to be on the inside (i) or outside (o) of the membrane. Phylogenetic analysis was performed in DNAstar's MegAlign program.

3.1.5 Results

3.1.5.1. Search for additional PXG genes in oat EST databases

The sequence of *AsPXG1*, the first peroxygenase gene identified in oat (Meesapyodsuk and Qiu, 2011), was used as a query sequence to search oat expressed sequence tag (EST) databases. Since *AsPXG1* was identified in an EST database prepared from developing seeds, oat EST databases prepared from developing seeds as well as from other tissues such as leaves, roots and germinating seeds were searched for homologous sequences. Assembly of the homologous EST sequences from the databases gave rise to six contig sequences which were named as Contig sequence 1 to 6, respectively. After the manual corrections of nucleotide inconsistencies in the sequence assembly, ORFs of each contig were determined as to the translational initiation and termination codons. The length of the consensus sequences of Contig 1 to Contig 6 including untranslated regions and ORF were 1014, 972, 940, 1200, 967 and 1173 bp long, respectively (Table 5). Among six contigs, Contig sequence 1, 5 and 6 contained full ORFs, while Contig sequence 2, 3 and 4 were partial, missing 3' ends of their ORFs.

Table 5. Sequence information of six *AsPXG* contig sequences

Contig sequence		ORF (bp)	Nucleotide identity with *AsPXG1*	PXG gene name
No.	Length			
1	1014	750	99.00%	*AsPXG1*
2	972	partial	56.20%	*AsPXG2*
3- 3.1	940	partial	50.60%	*AsPXG3*
3- 3.2	940	partial	48.20%	*AsPXG4*
4	1200	partial	39.00%	*AsPXG5*
5	967	654	36.00%	*AsPXG6*
6	1173	711	62.40%	*AsPXG7*

Contig sequence 1 showed 99% nucleotide identity with *AsPXG1* identified previously. The start codon was located at 205 nucleotides from the 5' end and the stop codon resided at 750 nucleotides downstream of the start codon, giving 205 bp long at the 5' untranslated region (UTR) and 59 bp at the 3' UTR. Due to its high nucleotide identity with the query, it was assumed to derive from the same gene *AsPXG1*. Contig sequence 2 to 6 showed nucleotide identity with *AsPXG1* in a range of 36 to 63% (Table 5). Contig sequence 2, 3 and 4 were missing 3' ends and the start codon was located respectively at 85, 101 and 89 bp from the 5' end. Contig 5 and 6 contained ORFs of 654 and 711 bp, respectively, with the start codon located at 230 bp and 292 bp from the 5' end, and the stop codon located at 84 and 170 bp from the 3' end, respectively.

Each consensus sequence was then used as the query sequence to search the NCBI (National Center for Biotechnology Information) EST databases from oat developing seeds of CDC Dancer in order to confirm the contig sequence in the cultivar. Accordingly, high homologous sequences were identified with each particular consensus sequence in the germplasm. The CDC Dancer sequence corresponding to each consensus sequence was then used as the template to design gene specific primers for retrieving the ORF for further analysis.

3.1.5.2. Isolation of putative full length PXG genes from oat

(a) Isolation of total RNA from oat developing seeds

Total RNA was isolated from CDC Dancer developing seeds using the TRIzol reagent. RNA integrity was assessed by running an aliquot of the RNA sample on a denaturing agarose gel stained with ethidium bromide. As seen in Figure 7, two bright bands representing two abundant RNA species, 28S and 18S ribosomal RNAs (rRNAs), in the total RNA was observed. As expected, the relative intensity ratio of the two bands was approximately 2:1. In addition, no sign of partially degraded RNA, i.e. smeared appearance of rRNA bands, was observed, indicating the RNA was of good quality. Subsequently the RNA was used as template for the biosynthesis of first-strand cDNAs used for RACE and ORF retrieval.

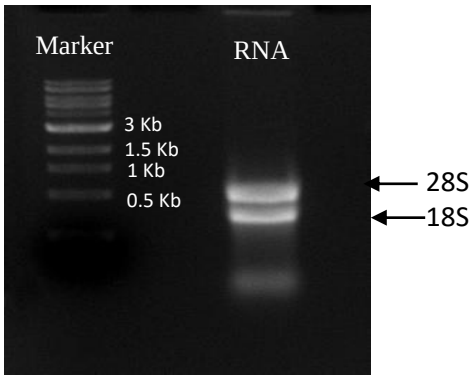


Figure 7. Agarose electrophoresis gel image of total RNA extracted from oat cultivar CDC Dancer developing seeds.

3.1.5.3 Retrieving missing ends of putative *AsPXG2* and *AsPXG3* genes

As in Table 5, Contig sequence 6 contained an ORF of 711 bp with about 63% nucleotide identity to the query. However, further sequence analysis showed that, unlike putative AsPXG2 and AsPXG3, this contig sequence (AsPXG7) did not align well with caleosin proteins at the N-terminal region. Therefore, Contig sequence 2 and 3 which showed both high sequence identity with the query and good alignment with caleosin proteins at the N-terminus region were chosen for further analysis.

As both Contig sequence 2 and 3 were not full length, missing 3’ ends, 3’ RACE PCR was thus performed to obtain the missing ends of the two sequences. The first-strand cDNA for the RACE was synthesized using a SuperScript III first-strand synthesis system. The RACE-PCR was performed in a series of two sequential PCR reactions. The initial PCR reaction was carried out using IB1 and IB3, two forward primers for the two sequences, and a common RACE outer primer to amplify the larger fragment with the expected size of about 500 bp in length.

After that, the expected fragment of the initial PCR product was eluted from the gel and used as the template for the nested PCR using IB2, IB4 and common RACE inner primers, which produced the product with the expected size at about 400 bps in length (Figure 8).

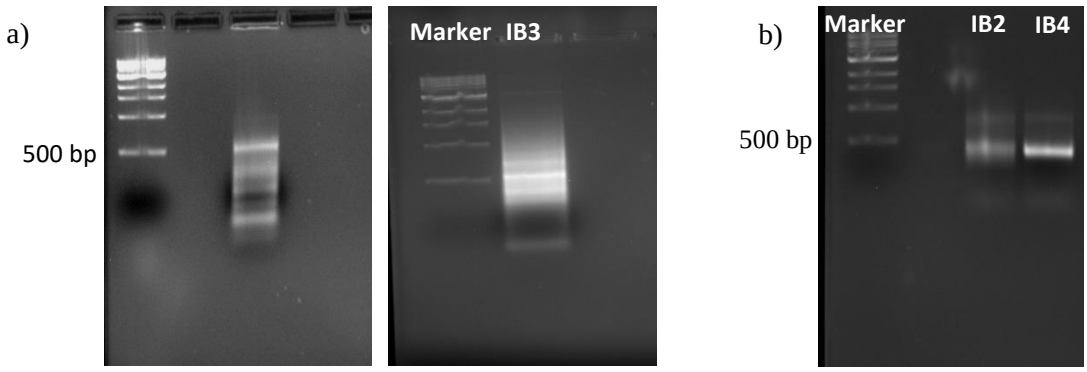


Figure 8. RACE–PCR for retrieving missing ends of putative *AsPXG* genes. a), Agarose electrophoresis gel images showing the first (outer) 3' RACE–PCR products of Sequence 2 and 3 respectively with the expected size (500 bp). b), A gel image of the second (nested) PCR products of Sequence 2 and 3 with the expected size (400 bp). IB1 and IB3 are the gene specific outer primers, and IB2 and IB4 are the gene specific inner primers for PCRs. Marker, 1 Kb ladder marker.

After the 3' nested RACE PCR reaction, PCR fragments with the expected size for Sequence 2 and 3 were isolated and ligated to a TA cloning vector pGEM-T and recombinant plasmids were transformed to *E. coli* TOP 10 electro-competent cells. Transformants carrying each nested PCR product were selected and the plasmids in the positive transformants were extracted and sequenced.

3.1.5.4 Sequence analysis of the putative AsPXG genes

Through the RACE, the missing 3' ends of the two sequences were successfully retrieved. The missing end of Sequence 2 was 208 bp in length. Assembly of this fragment with the other part of Sequence 2 revealed that a full-length ORF from start to stop codon was 702 bps long coding for 233 amino acids with predicted molecular mass of 26 kD. The start codon was located at 85 bps from the 5' end, and the stop codon was located at 105 bp from the 3' end (Figure 9). As for Sequence 3, the identified missing end was 185 bp in length and the stop codon was located at 88 bp from the 3'end. Assembly of this fragment with the other part of Sequence 3 revealed a full-length ORF of 627 bp long coding for 208 amino acids with predicted molecular mass of 26 kD (Figure 10). The two full length putative PXG genes were named as *AsPXG2* and *AsPXG3,* respectively.

The two ORF cDNAs including putative start and stop codons were then amplified by RT-PCR using gene specific primers and subsequently cloned into an expression vector and sequenced again prior to the expression study. The sequencing results of the full-length cDNAs retrieved by the RT-PCR of *AsPXG3* revealed an additional isoform sequence. Compared with the initial *AsPXG3* sequence, the isoform sequence exhibited 22 single nucleotide polymorphisms (SNP) throughout the sequence, which resulted in 11 amino acid changes in the translated protein (Figure 1 in Appendix). Therefore, the new isoform sequence was named as *AsPXG4* (Figure 2 in Appendix). Both *AsPXG3* and *AsPXG4* contained an ORF of 627 bp in length encoding a polypeptide of 208 amino acids. AsPXG4 shares 95% amino acid identity with AsPXG3. Further sequence comparison revealed that AsPXG2, AsPXG3 and AsPXG4 shared amino acid identity with AsPXG1 at about 51%, 44% and 42%, respectively (Figure 11). Alignment of these AsPXGs revealed that three histidine residues potentially for heme-binding and an EF-hand motif potentially for calcium-binding were highly conserved among these sequences. In addition, a proline knot motif with three conserved prolines over 12 residues believed to mediate a turn of an antiparallel α-helix or β-sheets was also somewhat conserved among the sequences (Figure 11). In addition, AsPXG2 has higher amino acid identity (65%) with a pollen caleosin from *Lillium longiforum*, while AsPXG3 and AsPXG4 have higher amino acid identity (77% and 74%) to a caleosin protein from *Triticum aestivum* and a peroxygenase-4 like enzyme from *Setaria italica* (71% and 69%).

Topology prediction of the three putative PXG proteins revealed a single transmembrane domain in all sequences, similar to the previously identified oat peroxygenase (AsPXG1). However, the location of the transmembrane domain in these proteins was slightly different (Figure 12). In AsPXG2, the predicted transmembrane domain was located between position 89 and 109, while in AsPXG3 and AsPXG4 it was located between position 54 and 74. As predicted, the regions prior to and behind the transmembrane domain in each sequence acquired polar structures residing on the membrane facing either the interior or exterior environment, respectively. According to the topology of these PXG proteins, AsPXG2 shared more close resemblance to the previously characterized AsPXG1, a Class I caleosin protein with the transmembrane domain approximately in the middle (Meesapyodusk and Qiu, 2011). On the other hand, AsPXG3 and AsPXG4 shared a similar topological structure with Arabidopsis AtPXG4, a Class II caleosin protein (Blee et al., 2012) where the transmembrane domain resided closely to the N-terminus.

Phylogenetic analysis of oat PXGs with related Arabidopsis PXGs showed that these PXGs could be grouped into two distinct clusters. The highly homologous AsPXG3 and AsPXG4 belonged to a cluster where Arabidopsis AtPXG4 was included, whereas AsPXG2 along with previously characterized oat AsPXG1 belonged to another cluster where two Arabidopsis AtPXG1 and AtPXG2 were included (Figure 13). This result is consistent with their topological structures.

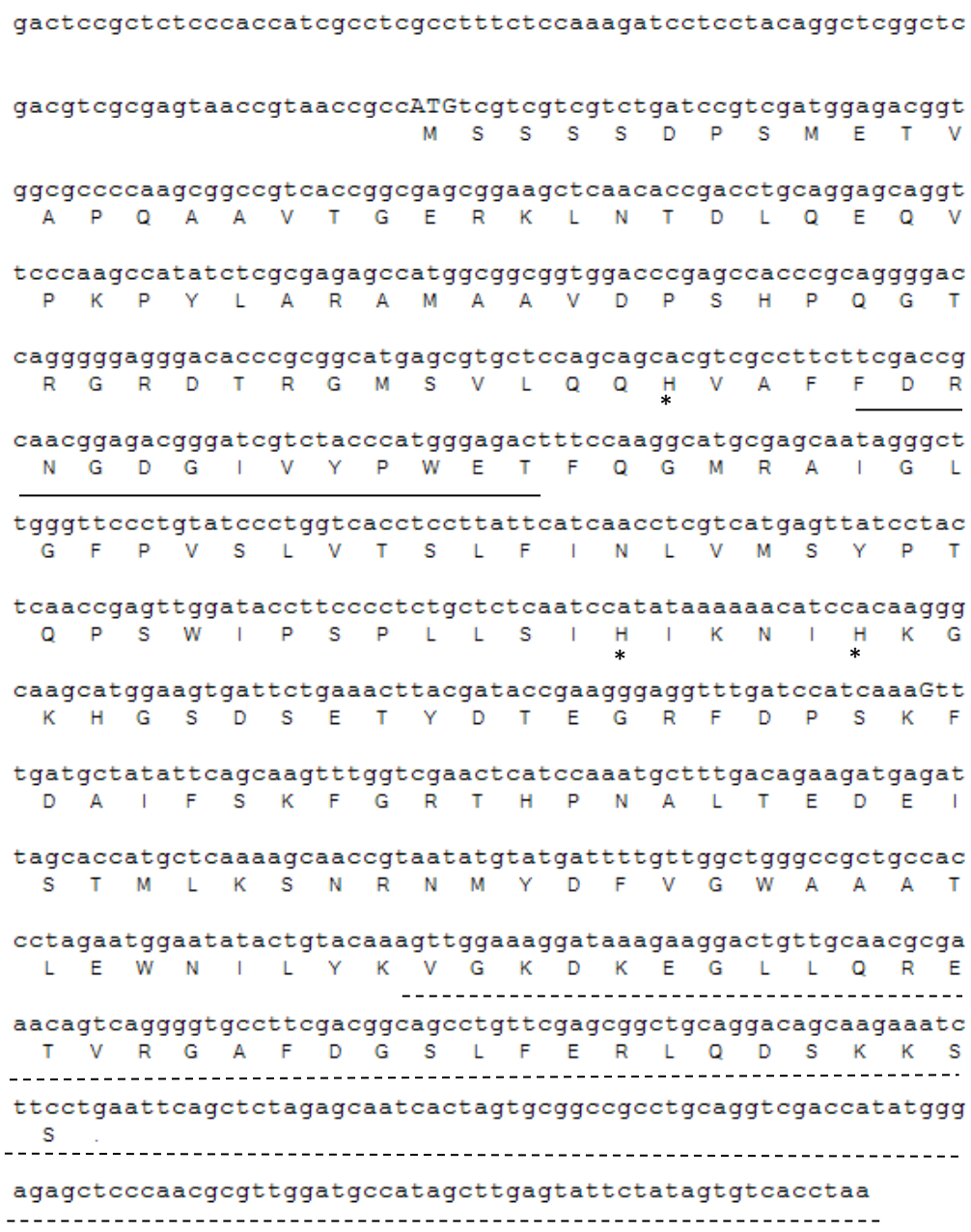


Figure 9. The full-length cDNA of *AsPXG2* and the translated open reading frame (ORF). The histidine residues for heme-binding and EF-hand motif involved in calcium-binding are highlighted with asterisks and a single solid line respectively. The sequence highlighted with a dashed line represents the sequence identified from 3'RACE PCR.

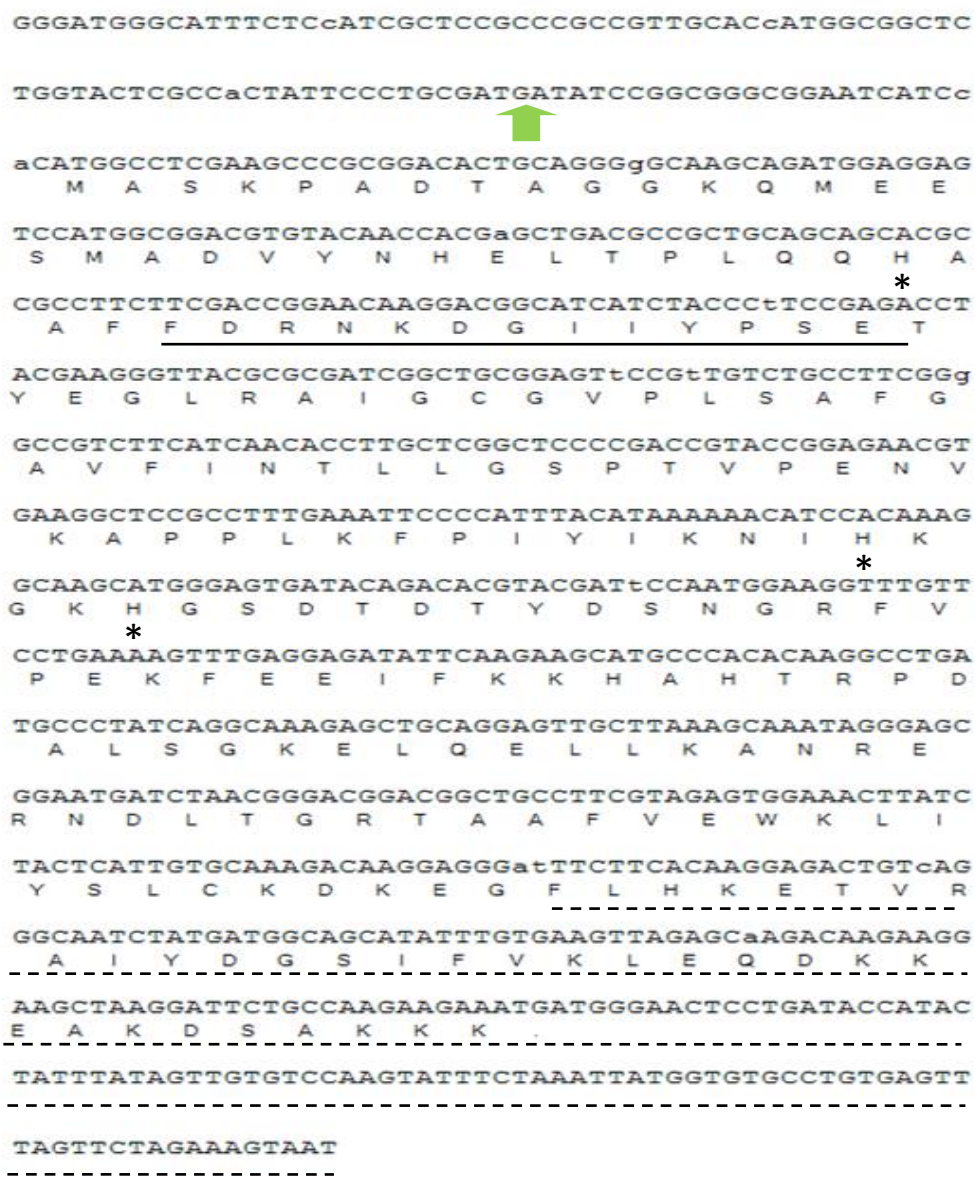


Figure 10. The full-length cDNA of *AsPXG3* and the translated open reading frame (ORF). The histidine residues for heme-binding and EF hand motif involved in calcium-binding are highlighted with asterisk and a solid line respectively. The sequence highlighted with a dashed line represents the sequence identified from 3'RACE PCR. The in-frame stop codon in the 5' upstream is indicated with an arrow.

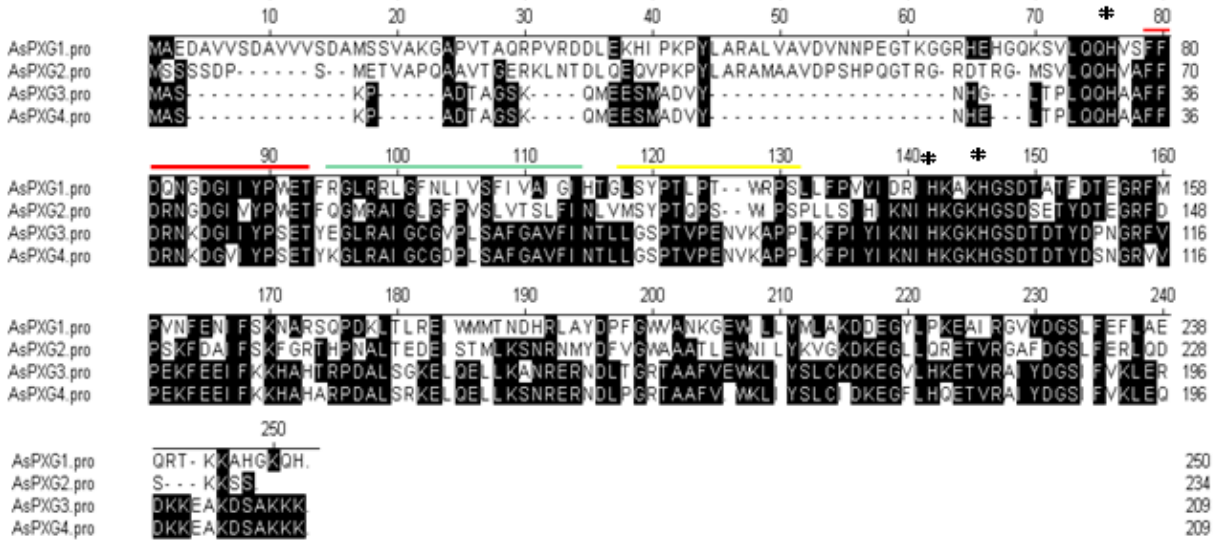
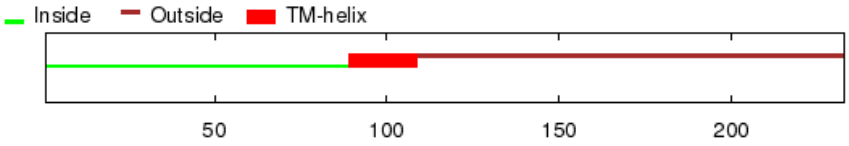


Figure 11. Sequence alignment of three putative PXG proteins AsPXG2, AsPXG3 and AsPXG4 with the query AsPXG1. The shaded regions indicate amino acids that are identical to each other. Conserved His residues, EF-hand motif and hydrophobic membrane-associated domain are highlighted with asterisks, red and green lines above the sequence, respectively. Moreover, proline knot motif is highlighted with a yellow single line. The amino acid number for the last residue in each row is listed on the right for each gene. Broken lines in the sequence represent gaps introduced for best alignment.

(A)



(B)

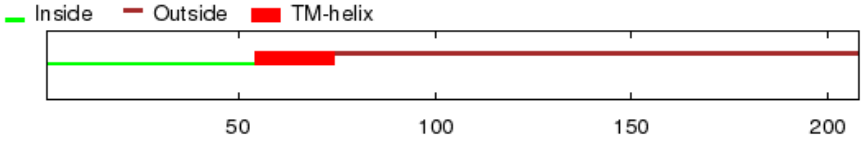


Figure 12. Predicted topology of putative oat PXGs by OCTOPUS topology predictor. Predicted topology of A is for AsPXG2 and B is for AsPXG3 and AsPXG4 amino acid sequences. Putative proteins are consisted with short transmembrane domain (20 amino acid residues) with globular loop regions facing interior and exterior of the organelle, respectively. Numbers in each panel represent amino acid number in each putative protein.

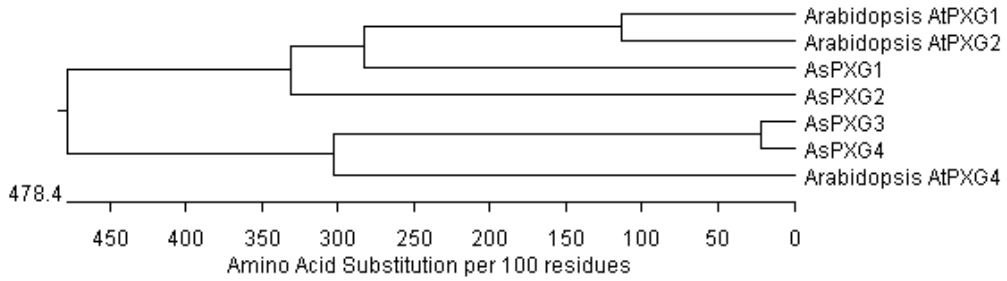


Figure 13: Phylogenetic analysis of oat PXGs with related Arabidopsis PXGs according to Jotun Hein aligning method. At=*Arabidopsis thaliana*; As=*Avena sativa*

3.1.6 Discussion

The previously identified oat peroxygenase (PXG) gene *AsPXG1* encodes a small protein with conserved heme-binding residues and a calcium-binding motif. It is solely expressed in developing seeds, not in other tissues such as roots, leaves, glumes and germinating seeds (Messapyodusk and Qiu, 2011). Like *AsPXG1*, three additional oat PXG genes, *AsPXG2*, *AsPXG3* and *AsPXG4*, identified in this study also encode small proteins with conserved heme-binding residues and the calcium-binding motif. Sequence analysis shows that these proteins belong to a family of caleosins, a group of small oil body-associated proteins with both calcium and heme binding ability (Hanano et al., 2006, Meesapyodusk and Qiu, 2011). Most of the caleosins identified so far have molecular masses around 25–29 kDa (Hanano et al., 2006). The predicted molecular weights of the three new PXG proteins fall in a similar range at about 27 kDa. All these data support that these newly-identified oat genes encode caleosin-like peroxygenases.

Naested et al., (2004) first classified caleosins into three groups based on their primary sequences and expression profiles (Naested et al., 2004). This classification was later updated by Hanano et al (2006) according to the location of the hydrophobic transmembrane domain, which anchors the protein into lipid bodies or endoplasmic reticulum (Hanano et al., 2006). Caleosins are characterized by the structure of three distinct domains: a hydrophobic central domain, a N-terminal region with a calcium binding EF-hand motif and a C-terminal region with kinase phosphorylation sites. Based on the location of the transmembrane domain in the primary structure, caleosin proteins are categorized into three classes. The first two classes mainly occur

in plants, while the third class occurs only in fungi. Class I caleosins have the hydrophobic transmembrane domain located in the middle of the protein, while the domain is shifted to the N-terminus in Class II caleosins. In Class III fungal caleosins, the domain appears randomly throughout the protein (Hanano et al., 2006). The genes encoding Class I caleosins have been well characterized. For instance, AsPXG1, a Class I claeosin-like PXG from oat is located with both lipid droplets (LD) and endoplasmic reticulum (ER) with high peroxygenase activity (Meesapyodusk and Qiu, 2011). However, little was known about Class II caleosins from plants until very recently. In 2012, a Class II caleosin protein AtPXG4 from Arabidopsis was characterized biochemically as a peroxygenase (Blee et al., 2012). As in AtPXG4, the transmembrane domain in AsPXG3 and AsPXG4 reside more closely to the N-terminus, whereas AsPXG2 has a middle transmembrane domain, showing its close resemblance to AsPXG1. In addition, phylogenetic analysis further reinforces this classification, showing AsPXG3 and AsPXG4 are clustered with Class II caleosin proteins such as AtPXG4, while AsPXG2 is clustered with Class I caleosin proteins such as AsPXG1 and AtPXG1.

Caleosins can reside in either lipid droplets or endoplasmic reticulum. Depending on the target location, caleosins might adopt different higher-dimensional structures. Caleosins, when binding to endoplasmic reticulum, were suggested to adopt a structure (Type I), in which the N-terminal is facing the lumen side while the C-terminus is facing the cytosol (Murphy, 2001). However, when residing in lipid droplets, caleosins would exhibit a different structure (Type II) where both their polar N- and C-terminal regions are exposed to the cytosol (Murphy, 2001) (Figure 14). All of the caleosins described to date including oat PXGs possess a single transmembrane domain, even though they belong to different classes of caleosins. It was hypothesized that the active site of AtPXG1 would face the lumen of the endoplasmic reticulum, whereas the phosphorelation sites would be in contact with the cytosol, adopting a Type I structure (Hanano et al., 2006). However, this assumption is not convincing, as the structure indicated that one conserved histidine residue from the N-terminus is located on the lumen side, while the other histidine residue(s) faces the cytosol. Such an orientation could obstruct the function of these histidine residues for heme co-ordination, which is essential for enzyme activity. As most of the oat peroxygenase activity is found in lipid droplets, we believe that the oat PXGs identified in this study would assume the Type II orientation regardless of their caleosin classes. In this structure, both N-terminus and C-terminus face the cytosol where

conserved histidine residues from both termini could readily interact to coordinate a heme group for the catalytic activity.

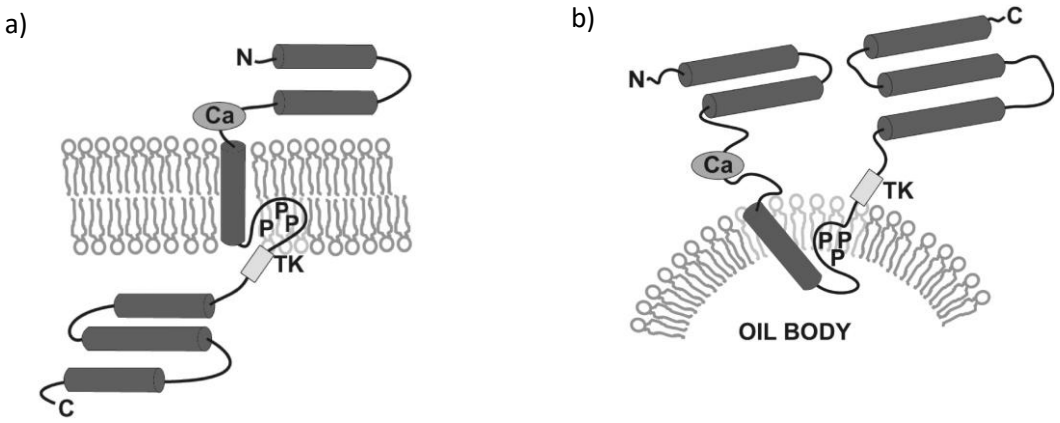


Figure 14. Approximate topological structure of caleosins. a) in the ER membrane and b) on the surface of a lipid body. N: N terminus, C: C-terminus, P-P-P: proline knot, TK: tyrosine kinase phosphorylation site, Ca: calcium-binding EF hand domain. According to Murphy, 2001.

AsPXG4 was retrieved using two specific primers for *AsPXG3*. The sequences of the two ORFs differ by only 22 nucleotides causing a change in 11 amino acids in the translated proteins. Although both sequences were retrieved by the same pair of primers, they might derive from different PXG loci rather than the same locus with alternatively splicing transcripts, as alternatively splicing products generally possess indel sequences, which is not observed between the two sequences. The oat cultivar CDC Dancer is a hexaploid. The polyploidy nature of the genome implies duplication and rearrangement of the primitive genomes may have occurred many times during evolution. Thus, the current oat genome may contain many copies of PXG genes that are highly homologous with each other, such as AsPXG3 and AsPXG4. Although two AsPXGs differ in 11 amino acids, these amino acid substitutions do not occur on or adjacent to the predicted active sites involved in heme binding.

3.2 Study 2

Functional analysis of oat putative *AsPXG* genes

3.2.1 Introduction

Oat has been known to produce several epoxy fatty acids by peroxygenase (PXG) that account for up to 5.0% of the total fatty acids in seeds. These include 9,10-epoxy-18:0, 9,10-epoxy-18:1-12*c* and 12,13-epoxy-18:1-9*c* and are mainly present in neutral lipids such as triacylglycerols and diacylglycerols (Leonova et al., 2008: Doehlert et al., 2010; Meesapyodsuk and Qiu, 2011). Due to this nature, oat serves as a model system for analyzing the mechanism underlying the biosynthesis of epoxy fatty acids through the peroxygenase pathway (Hamberg and Hamberg, 1996). The previous study in this thesis describes cloning three putative peroxygenase genes from oat. This study will examine their function by expressing them in microbial systems for *in vitro* assays.

3.2.2 Abstract

Functional analysis of three genes encoding putative PXG from oat was carried out in *E. coli*. The coding regions of individual genes were cloned into pET28a behind the T7 promoter. Lysates of transformants expressing the genes were used as enzyme sources for *in vitro* assays. GC analysis of the *in vitro* assay products showed that transformants of *AsPXG3* and *AsPXG4*, like that of the positive control *AsPXG1*, produced 9,10-epoxystearic acid methyl ester in the presence of oleic acid methyl ester, indicating they encode a functional peroxygenase. *AsPXG3* had high specific activity at 42 µmol/mg/min with a substrate conversion efficiency at 25%, whereas the activity of *AsPXG4* was slightly lower at 38 µmol/mg/min with 21% substrate conversion efficiency. AsPXG2 did not show any activity in *E. coli*, but showed low activity in *Pichia pastoris*. Substrate specificity of AsPXG3 was determined using a variety of free unsaturated fatty acids as substrates in *in vitro* assays. The result showed it possessed activity towards a broad range of substrates from medium chain to very long chain unsaturated fatty acids with one to three double bonds. Among these fatty acids, 9,12,15-*cis*-octadecatrienoic acid (18:3-9*c*,12*c*,15*c*, linolenic acid) was the most preferred substrate. These results indicate that AsPXG3, a Type II caleosin from oat, could effectively epoxidize unsaturated fatty acids that

commonly occur in plants with the long chain polyunsaturated fatty acid as the preferred substrate.

3.2.3 Hypothesis

Putative oat peroxygenase genes *AsPXG2, AsPXG3* and *AsPXG4* encode functional peroxygenase producing corresponding epoxy fatty acids in the presence of unsaturated fatty acid substrates.

3.2.4 Experimental Approach

3.2.4.1 Heterologous expression of AsPXGs

To determine the function of *AsPXG2, AsPXG3* and *AsPXG4,* ORFs of the three genes were cloned into an *E. coli* expression vector pET28a under guidance of the T7 promoter (Figure 15). The pET system is specifically designed for cloning and expressing recombinant proteins in *E. coli*. The translational vector possesses a highly efficient ribosome binding site from the phage T7 major capsid protein gene and has reading frame compatibility with any heterologous gene (Novagen pET system manual, 2010).

When cloning the genes of interest (GOIs) into the expression vector, the following protocols for insert and vector preparation and ligation were used. For insert preparation, 1 µg of GOI-DNA, 1 µL of 10 U/µL *Eco*RI restriction enzyme and 5 µL of corresponding 10X buffer were mixed with appropriate volume of water. The reaction was incubated for 2 hours at 37°C. For vector preparation, pET28a was digested with the same restriction enzyme in another reaction. After the digestion, 1 µL Calf Intestinal Alkaline Phosphatase (CIAP) was added to the digested mixture and incubated for 10 minutes at 37°C to dephosphorylate the 5' terminus in order to eliminate self-ligation prior to insert/vector ligation. The digested products were separated and purified from an agarose electrophoresis gel. The purified fragments of GOI and pET28a were mixed in a 3:1 ratio in a reaction with 2 µL 5X buffer, 0.5 µL T4 ligase, and 7.5 µL water. The reaction was incubated at 16°C overnight and the recombinant plasmids were then transformed into the intermediate/nonexpression host *E. coli* Top10. Transformants were selected on a kanamycin medium (50 µg/mL) and screened by PCR using a primer in the promoter i.e. T7, and a primer specific for the gene. Once confirmed in the non-expression host, the recombinant plasmid was extracted and transformed into an expression host strain

Rosetta2(DE3)pLysS bearing the T7 RNA polymerase gene for expression of target genes. The transformants were selected on kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL) medium. All constructs were checked by sequencing prior to the expression in the Rosetta2(DE3)pLysS strain. The enzyme assay was conducted essentially as in Meesapyodsuk and Qiu (2011). Rosetta2(DE3)pLysS/pDM70 was used as a positive control as pDM70 carries the previously characterized *AsPXG1*.

Each *E.coli* transformants carrying putative *AsPXG2, AsPXG3* and *AsPXG4* together with the positive control and empty vector were first grown at 37 ºC overnight in 10 mL Luria-Bertani medium (LB) containing kanamycin (50 µg/mL) or ampicillin (100 µg/mL) along with chloramphenicol (34 µg/mL).

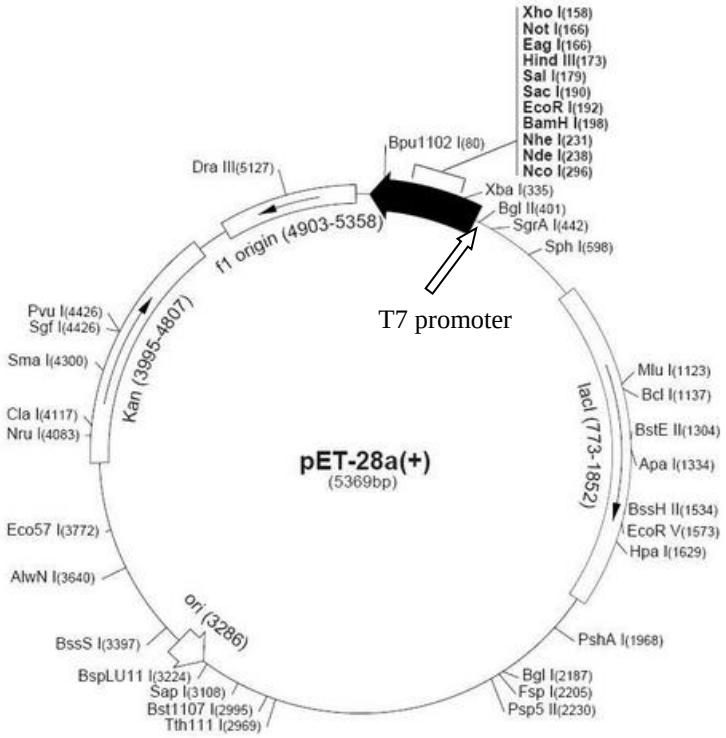


Figure 15. pET28a vector map. This expression system was used for functional analysis of putative PXG genes. The T7 promoter is located at 370-386 bp in the vector backbone. (From Novagen, 2010)

The overnight cultures were then inoculated into 10 mL of fresh LB medium containing the same antibiotics. When the culture was grown at 37ºC until an OD_{600} = 0.5 to 1.0, IPTG was added to a final concentration of 1 mM to induce the expression at 30 ºC for 4 hours. After induction, the cells were harvested by centrifugation (5000 rpm, 15 minutes). The cell pellets

were washed twice with 50mM Tris-HCl (pH 7.5, 5000 rpm, and 15 minutes) and resuspended in a homogenizing buffer containing 50 mM Tris-HCl (pH 7.5), 0.6 M sorbitol, 10% glycerol, 0.2% emulphogene. The transformant cells were disrupted using a bead beater. The disrupted samples were centrifuged at 4 ^{o}C (8000 rpm, 10 minutes) and the supernatant was used as the enzyme source in the *in vitro* assay. Protein concentrations in each sample were determined according to Bradford method using bovine serum albumin as the standard (Bradford, 1976).

To functionally characterize the putative peroxygenase *AsPXG2*, the coding region was cloned into the yeast expression vector pPICZ-B at the *Eco*RI restriction site and under the control of the AOX1 promoter following the restriction digestion described above. Digested products were ligated in a 3:1 ratio and the constructs were transformed into the intermediate host Top10 *E. coli*. Transformants were selected on low salt LB plates containing zeocin (25 µg/mL) and screened by PCR using a primer in the promoter i.e., 5' AOX1 and a gene specific primer. To promote integration into the yeast genome, the constructs carrying GOI were linearized with *Dra*I. Once the complete linearization was confirmed by agarose gel electrophoresis of a small aliquot of the digest, the linearized vectors were extracted once with phenol/chloroform. In this process, 200 µL of phenol/chloroform/isoamyl alchohol were added to the linearized vector. After vortexing, the sample was centrifuged at 12000 rpm for 5 min. The top phase was collected and transferred to a new tube and the volume was adjusted to 200 µL with water. Following that, 20 µL of 3M sodium acetate, 450 µL of 100% EtOH was added and the reaction was incubated for 30 min at -80 ^{o}C. Then, the mix was centrifuged at 12000 rpm for 5 min and pelleted DNA was washed with 1 ml of 80% EtOH. After that, the DNA pellet was air dried and dissolved in 10 µL of sterile water and stored at -20 ^{o}C.

Pichia pastoris X-33 strain, a highly successful system for the production of a wide variety of recombinant proteins was prepared for electroporation as follows. 50 ml of fresh yeast extract peptone dextrose medium (YPD) was inoculated with 20 µL of *Pichia pastoris* culture grown overnight at 28 ^{o}C and re-cultured overnight again. Once the OD_{600} = 1.3-1.5, the cell pellet was collected after centrifugation at 1500×g for 5 minutes at 4°C. The cell pellet was re-suspended in 50 mL of ice-cold (0–4°C) sterile water. Then three centrifugations and resuspentions were performed with 25 mL of ice-cold (0–4°C) sterile water and 2 mL of ice cold (0–4°C), 1 M sorbitol respectively. Finally, the cell pellet was resuspended in 100 µL of ice cold 1M sorbitol and used for the transformation.

For the transformation, 100 µL of the cells prepared above were mixed with 5–10 µg of the linearized construct pPICZ/AsPXG2, which was then transferred to an ice-cold (0–4°C) 0.2 cm electroporation cuvette. After incubation for 5 minutes on ice, the mixture was pulsed at 2000 volts using an electroporator. Immediately after, 1 mL of ice-cold 1 M sorbitol was added to the cuvette, which was then transferred to a sterile 15 mL tube and incubated at 30°C without shaking for 1 to 2 hours. After that, 50-200 µL of the mixture was sprayed on YPD plates containing zeocin (100 µg/mL). The plates were incubated for 2–3 days at 30°C. Transformants were screened by PCR using a primer in the promoter i.e., 5' AOX1 and a primer specific for the gene.

For the expression study, the transformed cells were grown at 28 ^{o}C for 24 hours with shaking in a buffered minimal glycerol medium (BGM). The culture was centrifuged and resuspended in a buffered minimal methanol medium (BMM). After 60 hours, the resulting culture was harvested and washed once with a Tris buffer (50 mM Tris-HCl pH 7.5). The pellet was resuspended in a homogenizing buffer containing 50 mM Tris-HCl (pH 7.5) and 0.6 M sorbitol, and the cells were disrupted with glass beads using a Mini-Beadbeater (Biospect; three cycles of 1 minute each). After the centrifugation, the supernatant was used as the enzyme source for the *in vitro* assay. The yeast transformant X-33/pDM71carrying *AsPXG1* was used as the positive control in the *in vitro* assay (Meesapyodusk and Qiu, 2011).

3.2.4.2 *In vitro* enzyme assay

Peroxygenase assays were carried out in a volume of 200 µL containing 0.1 mg protein, 0.5 mM cumene hydroperoxide, 0.5 mM oleic acid methyl ester (ME) in Buffer D containing 100 mM Tris-HCl (pH 7.0), 20% glycerol and 0.4% emulphogene. For quantifying the products, 3 µg of 17:0-ME was used as an internal standard. The reaction was incubated at 45 ^{o}C for 15 min with shaking and then stopped by adding 500 µL ethyl acetate containing 10 µL of glacial acetic acid. The fatty acid methyl ester (FAME) products were extracted twice with 500 µL ethyl acetate and dried under the N_2 gas. Dried FAMEs were dissolved in 50 µL of ethyl acetate and analyzed on an Agilent 6890N gas chromatography (GC) equipped with a DB-23 column (30-m x 0.25-mm) with 0.25 µm film thickness (J&W Scientific). Initially the column temperature was maintained at 160 °C for 1 minute, and then raised to 240 °C at a rate of 4 °C /minute. Identity of the products was confirmed by comparing GC retention time with those of standard fatty acids,

such as C14:0, C14:1, C15:0, C16:0, C16:1-9*c*, C17:0, C18:0, C18:1-9*c*, C18:1-11*c*, C18:2-9*c*,12*c*, C18:3-6*c*,9*c*,12*c*, C18:3-9*c*,12*c*,15*c*, C20:0, C20:1-5*c*, C20:1-8*c*, C20:1-11*c*, C20:2-11*c*,14*c*, C22:0, C22:1-13*c*, C24:0, C24:1-15*c*, and C26:0.

3.2.4.3 Substrate specificity of peroxygenase

In order to determine the substrate specificity of AsPXG3, a variety of free fatty acids (FFA) such, as mono and poly unsaturated fatty acids ranging from C14 to C24, were provided as the substrates in the *in vitro* assays (Table 6). For quantification of the fatty acid products, 3 µg of C17:0 FFA was used as the internal standard. As mentioned before the *in vitro* assays were carried out at 45 ^{o}C for 15 minutes with 500 rpm shaking in a volume of 200 µL using cumene hydroperoxide as oxidant. After the reaction, the free fatty acid products were methylated using 25 µL of methanol and 50 µL of diazomethane. The fatty acid methyl ester (FAME) products were dried again under N_2 gas and then dissolved in 50 µL of ethyl acetate and analyzed by gas chromatography (GC).

3.2.5 Results

3.2.5.1 Functional analysis of the putative peroxygenase genes by *in vitro* assays

To functionally analyze the three putative PXG genes cloned from oat, the ORFs were cloned into an *E. coli* expression vector separately. The activity was determined by *in vitro* assays using lysates from transformants expressing the genes. As shown in Figure 16, in the presence of cumene hydroperoxide as an oxygen donor and oleic acid-methyl ester as a substrate, both transformants of *AsPXG3* and *AsPXG4* produced a new peak compared to the negative control. This peak had the same chromatographic retention time as the standard 9,10-epoxystearic acid. This result showed that transformants of *AsPXG3* and *AsPXG4*, like that of the positive control *AsPXG1*, produced a corresponding epoxy fatty acid (9,10-epoxy-18:0) in the presence of an unsaturated fatty acid substrate (18:1-9*c*), indicating they all encode functional peroxygenase (Figure 16).

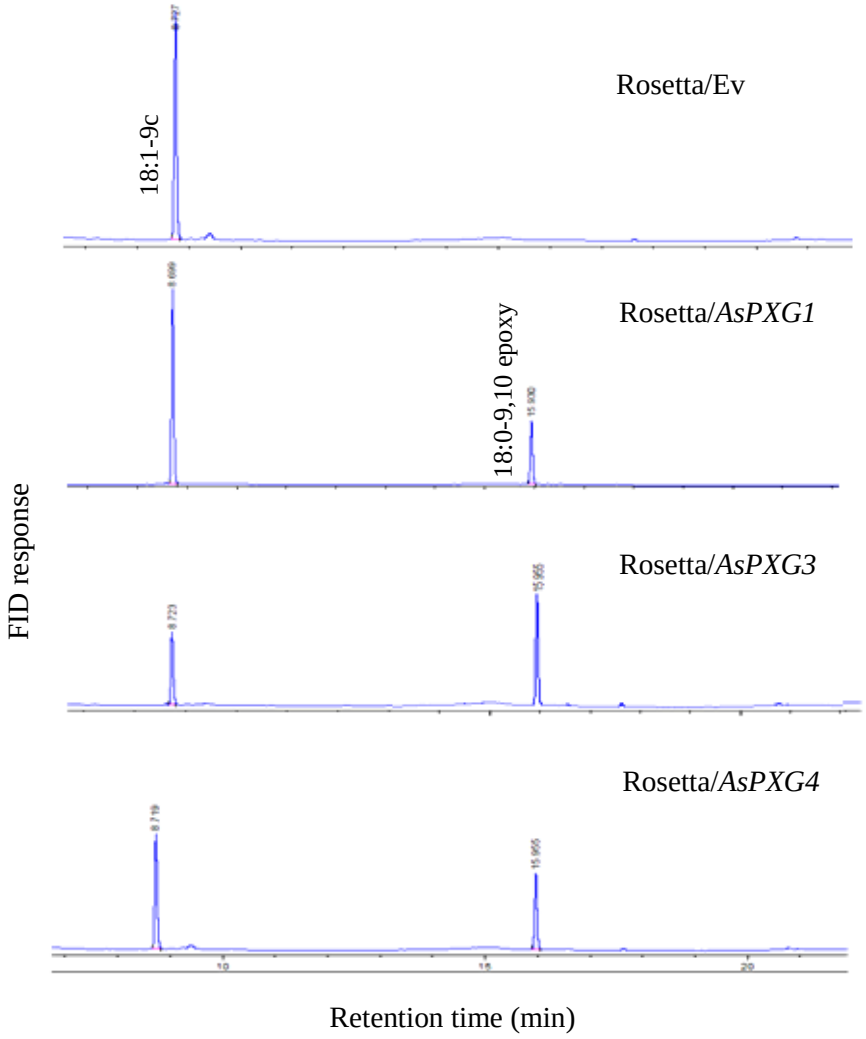


Figure 16. Functional characterization of oat putative peroxygenase *AsPXG3* and *AsPXG4* in *E. coli* Rosetta2(DE3)pLysS strain in the presence of cumene hydroperoxide as an oxygen donor and oleic acid methyl ester as a substrate. The four panels show fatty acid methyl ester products prepared from enzymatic reactions using extracts of Rosetta/EV (negative control), Rosetta/*AsPXG1* (positive control), Rosetta/*AsPXG3* and Rosetta/*AsPXG4* respectively.

The positive control AsPXG1 had specific activity at 35 µmol/mg/min and about 18% of substrate conversion efficiency. However, *AsPXG3* had higher specific activity 42 µmol/mg/min and substrate conversion efficiency (about 25%) than the positive control AsPXG1. *AsPXG4* had activity at 38 µmol/mg/min and about 21% conversion efficiency, which was also slightly higher than the positive control (Figure 16).

AsPXG2 did not show any activity when expressed in *E. coli* (data not shown). Subsequently, this putative gene was cloned into pPICZ-B, a yeast expression vector for the expression in *Pichia pastori,* as the yeast was believed to be a better host for expressing plant peroxygenase (Meesapyodsuk and Qiu 2011). GC analysis of the *in vitro* assay products showed that yeast transformants of *AsPXG2,* like that of the positive control *AsPXG1* (pDM 71), produced 9,10-epoxystearic acid methyl ester (9,10-epoxy-18:0), although at a low level, in the presence of C18:1-9 methyl ester, indicating it could function as peroxygenase, but with low activity. The specific activity of AsPXG2 in yeast was at 3 µmol/mg/min with only 2 % substrate conversion efficiency (Figure 17).

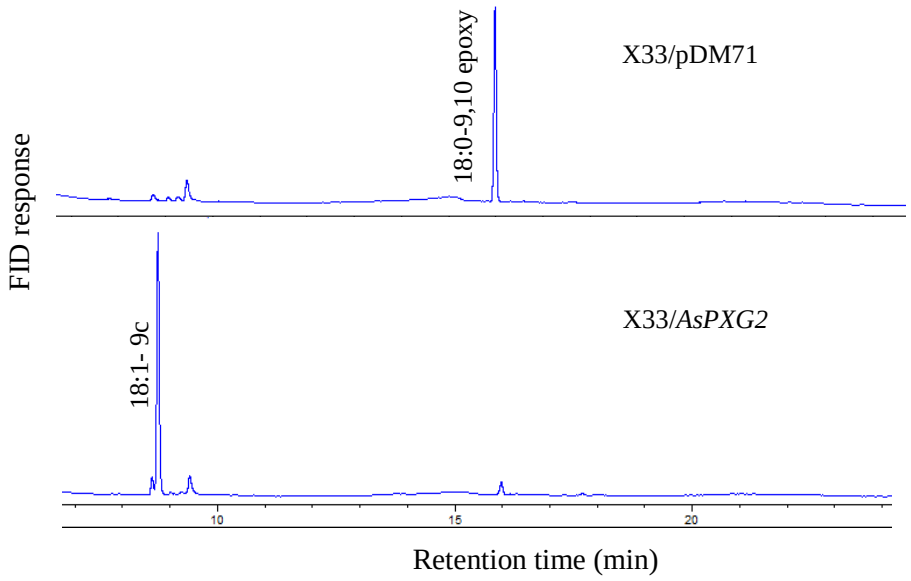


Figure 17. Functional characterization of oat putative peroxygenase *AsPXG2* in *Pichia passtoris* X-33 strain in the presence of cumene hydroperoxide as an oxygen donor and oleic acid methyl ester as substrate. The two panels show fatty acid methyl ester products from enzymatic reaction extracts of X-33/pDM71 (positive control) and X-33/*AsPXG2,* respectively.

3.2.5.2 Substrate specificity of oat peroxygenase AsPXG3

As seen above, three new oat PXGs displayed significant differences in enzymatic activity. AsPXG2 did not show any activity when expressed in *E. coli*, though very low activity was detected when expressed in yeast. On the other hand, both AsPXG3 and AsPXG4 (95% amino acid identity with each other) showed high PXG activity in *E. coli*. Between AsPXG3 and

AsPXG4, AsPXG3 exhibited higher enzymatic activity than AsPXG4. Therefore, substrate specificity of AsPXG3 was further analyzed by the *in vitro* assay.

To determine the substrate specificity of AsPXG3, a variety of free fatty acids ranging from 14 to 24 carbon chain lengths with one to three double bonds were used for the *in vitro* assays. As shown in Table 6, AsPXG3 exhibited enzymatic activity towards all fatty acids tested in the *in vitro* assays, though the specific activity differed considerably among these substrates. Noticeably, AsPXG3 possessed significantly higher specific activity towards long-chain polyunsaturated fatty acids such as linolenic acid (C18:3-9*c*,12*c*,15*c*) and linoleic acid (C18:2-9*c*,12*c*).

Table 6: The substrate preference of AsPXG3. Specific activity of AsPXG3 was determined by the *in vitro* assays with different substrates.

Substrate	Specific Activity (μmol/min/mg) Mean ± SD
C14:1-9*c*	167 ± 17
C16:1-9*c*	240 ± 04
C18:1-9*c*	208 ± 03
C20:1-11*c*	203 ± 05
C22:1-13*c*	143 ± 11
C24:1-15*c*	144 ± 07
C18:2-9*c*,12*c*	685 ± 17
C18:3-9*c*,12*c*,15*c*	746 ± 22

Number of replicates in each treatment = 3

On average, the activity on the polyunsaturated fatty acids was three to four times that on monounsaturated fatty acids. Between two polyunsaturated fatty acids examined, higher activity was detected on linolenic acid over linoleic acid, indicating linolenic acid is the more preferred substrate for AsPXG3 under the assay condition. Among monounsaturated fatty acids, less activity was detected on very long-chain mono-unsaturated fatty acids such as *cis*-15-tetracosenoic acid (C24:1-15*c*) and 13-*cis*-docosenoic acid (C22:1-13*c*), while higher activity

was observed on long chain monounsaturated fatty acids such as 9-cis-hexadecenoic (C16:1-9*c*) and 9-cis-octadecenoic (C18:1-9*c*, oleic acid).

3.2.6 Discussion

In this study, three new oat PXG genes from oat *AsPXG2*, *AsPXG3* and *AsPXG4* were functionally characterized. When expressed in *E. coli*/yeast, they produced 9,10-epoxystearate (9,10-epoxy-18:0) in the presence of C18:1-9*c*, indicating these genes encode functional peroxygenase catalyzing hydroperoxide dependent monooxygenation of unsaturated fatty acids producing corresponding epoxy fatty acids.

The previously characterized oat AsPXG1 is a membrane associated enzyme located with both lipid particle and endoplasmic reticulum. It can be efficiently solubilised from the membrane by detergent, as the membrane-association is weak with a single trans-membrane domain (Meesapyodusk and Qiu, 2011). When expressed in yeast or *E. coli*, AsPXG1 catalyzes strictly hydroperoxide dependent epoxidation of unsaturated fatty acids. It can only use free or methylated unsaturated fatty acids with double bonds in *cis* configuration as substrates and the most preferred substrate for AsPXG1 is oleic acid (Meesapyodusk and Qiu, 2011). In the present study, the peroxygenase activity of three new oat genes was examined using oleic acid-methyl ester as a substrate and the substrate specificity of AsPXG3, the highest active gene among the three, was determined using a variety of free unsaturated fatty acids. Like AsPXG1, AsPXG3 also possesses activity on broad substrates ranging from medium chain to very long chain unsaturated fatty acids. Similarly, AsPXG3 could use both fatty acid methyl ester and free fatty acids as substrates, however, the specific activity of AsPXG3 on free oleic acid was approximately six times that on fatty acid methyl ester. This result suggests that free fatty acid is likely the biological substrate for oat PXG3.

Arabidopsis PXG4 is a Class II caleosin that can efficiently catalyze hydroperoxide-dependent fatty acid epoxygenation reactions when expressed in yeast (Blee et al., 2012). In contrast to oat AsPXG1, a Class I caleosin with oleic acid being the most preferred substrate (Meesapyodusk and Qiu, 2011), AtPXG4 shows higher catalytic efficiency with polyunsaturated fatty acids such as linoleic (C18:2-9*c*,12*c*) and linolenic acids (C18:3-9*c*,12*c*,15*c*). These fatty acids are readily oxidized into their corresponding monoepoxides and diepoxides *in vitro*. Similar to AtPXG4, AsPXG3 also shows much higher activity on two long chain

polyunsaturated fatty acids, linoleic and linolenic acid, over monounsaturated fatty acids tested. This result further confirms that AsPXG3 is a Class II caleosin from oats (*Avena sativa*).

According to GC analysis, the transformants with *AsPXG3* produced only one major new peak compared to the negative control when two polyunsaturated fatty acids were used as the substrates in the *in vitro* assays. At this stage no further analysis was performed for identity clarification of the epoxy fatty acid products formed during the *in vitro* assay. However, previous results show that incubation of linoleic and linolenic acids with AtPXG4 in the presence of cumene hydroperoxide produces two groups of new compounds *in vitro* namely monoepoxides and diepoxides. Further, the rate of formation of monoepoxides from linoleic acid by AtPXG4 is about two orders of magnitude faster than the formation of diepoxides. It possesses high and constant preference for the formation of monoepoxide 9,10-epoxy-12-octadecenoic acid during the linoleic acid epoxidation. However, AtPXG4 appears far more pronounced in diepoxide formation from linolenic acid. The oxidation of a *cis* double bond at position C9,10 prevails over that at positions C12,13 and C15,16 when linolenic acid was used as the substrate. Based on these data, it can be assumed that similar to AtPXG4, epoxidation of polyunsaturated fatty acids by AsPXG3, a Class II caleosin, should take place preferentially at position C9,10, then at C15,16 and lastly at C12,13, which is in contrast to AsPXG1 which mostly prefers epoxidation of the C12,13 double bond of linoleic acid (Meesapydusk and Qiu, 2011: Blee et al., 2012). Accordingly the major product when two polyunsaturated fatty acids are used as substrates in the assays is most likely monoepoxide 9,10-epoxy-12-octadecenoic acid from linoleic acid or diepoxide 9,10-15,16-diepoxy-12-octadecenoic acid from linolenic acid.

Interestingly, we could not observe any peroxygenase activity of AsPXG2 when it was expressed in *E. coli*, although it has high sequence similarity to AsPXG1 that is highly active in *E. coli*. The reason for this is various. First or most likely, it might be an authentic caleosin with naturally very low peroxygenase activity because of the structure. Secondly, *E. coli* might not be a good system to express this specific gene. Indeed, low activity was detected when it was expressed in yeast *Pichia pastoris.* However, the activity is still very low compared with the control in the yeast. Regardless, this result indicates that the yeast is a better host for expressing the oat gene, which is likely due to the superior expression of genes from eukaryotes such as oat by more optimal regulation at one or multiple levels such as transcriptionally, translationally and post-translationally.

Study 3

Structure-function analysis of oat peroxygenase through site-directed mutagenesis

3.3.1 Introduction

Peroxygenase (PXG) is a hydroperoxide-dependent oxygenase that catalyzes epoxide formation on unsaturated double bonds in fatty acids (Piazza et al., 2003; Meesapyodusk and Qiu, 2011). Previous studies showed peroxygenase belongs to the caleosin family where histidine residues for heme-binding are critical for the enzyme activity (Piazza et al., 2003; Hanano et al., 2006). Sequence alignment of plant caleosins shows that a number of histidines are conserved among sequences. Site-directed mutagenesis on Arabidopsis PXG1 revealed that the catalytic activity is dependent on two highly conserved histidines (Hanano et al., 2006). However, no studies have been undertaken to examine functions of other conserved residues in any PXG enzymes. The two previous studies in this thesis describe cloning and functional analysis of peroxygenase genes from oats. This study attempts to interrogate functions of all the conserved histidines as well as the conserved residues surrounding these histidines of an oat PXG, AsPXG3, by site-directed mutagenesis. Elucidation of the enzymatic properties such as specific activity and substrate specificity of the mutants would facilitate our understanding on the structure-function relationship of the peroxygense.

3.3.2 Abstract

Sequence alignment of oat AsPXG1, AsPXG2, AsPXG3 and AsPXG4, as well as *Arabidopsis* AtPXG1 and AtPXG2, revealed that three histidine residues were fully conserved in three regions of the proteins. To study the structure-function relationship of AsPXG3 peroxygenase, the three histidines and nine conserved residues surrounding these histidines were individually replaced by alanine using overlapping extension PCR technique. Each mutant with a single amino acid substitution was heterologously expressed in *E. coli* and its PXG activity was determined by *in vitro* assays. The site-directed mutagenesis analysis showed that substitution of the first conserved histidine at position 32 (H1) and the third histidine at position 102 (H3) of AsPXG3 with alanine respectively resulted in complete loss of the enzymatic activity, while substitution of the second conserved histidine residue at position 98 (H2) resulted in only slight reduction (~25%) of the PXG activity. This result indicates that only H1 and H3 are absolutely

essential for the peroxygenase. Substitution of leucine at position 29, isoleucine at position 97, and lysine at position 101 with alanine reduced the enzymatic activity by more than 80% compared to the wild type, indicating these three residues are also very important for the catalytic activity. Substitution of glutamine at position 30, serine at position 104 and aspartic acid at position 105 resulted in reduction of the activity by 50% to 60%, indicating these residues have some roles for the enzymatic activity. However, the mutants with substitution of glutamine at position 31, lysine at position 99 and glycine at position 103 respectively retained more than 75 % of the original activity, implying that these residues might not have significant roles in the PXG activity. For better understanding on the structure-function relationship of the PXG, substrate specificity of the three mutants where leucine (M1), isoleucine (M5) and lysine (M8) were individually substituted was determined using a variety of free fatty acids as the substrates in the *in vitro* assays. Compared with the wild type enzyme, the activities of the three mutants were drastically reduced on all the mono-unsaturated fatty acids as well as linoleic acid (C18:2-9,12). However, the three mutants exhibited activity on linolenic acid (C18:3-9,12,15) very differently. Particularly, isoleucine replacement (M5) led to only slight reduction of the activity (about 15%) and lysine replacement (M8) resulted in an increase in the activity by 12%. This result suggests that these conserved residues might play roles in defining the shape and size of the catalytic pocket for interaction of the heme with fatty acid substrates.

3.3.3 Hypothesis

Site-directed mutagenesis by overlapping-PCR is a powerful tool for introducing specific and intentional changes to a DNA-coding sequence resulting in a mutated protein. This approach can be used to examine function of histidines and conserved amino acid residues in an oat peroxygenase.

3.3.4 Experimental approach

3.3.4.1 Sequence alignment and mutation site determination

Amino acid sequence of *AsPXG1*, the first peroxygenase gene identified in oat, was used as the query sequence (Meesapyodsuk and Qiu, 2011) to search other PXG homologs from oat EST databases that are publically available from the National Center for Biotechnology Information (NCBI). The search resulted in the identification of three new EST sequences

(AsPXG2, AsPXG3, and AsPXG4) which were highly homologous but not identical to AsPXG1. Three new AsPXGs along with AsPXG1 and AtPXG from *Arabidopsis* were aligned using the DNAstar program (Lasergene 9 Core Suite, Madison, USA) to identify highly conserved regions that might be important for PXG activity. Three highly conserved regions each with a histidine and a few surrounding residues were identified as the target sites for the site-directed mutagenesis.

3.3.4.2 Site-directed mutagenesis and AsPXG3-mutants development

In order to investigate the role of the selected conserved amino acid residues, twelve mutants were constructed in which each of these residues was replaced by alanine. Alanine was the substitution residue of choice, as it eliminates the effect of the side chain beyond the β-carbon and yet does not alter the main-chain conformation as some residues such as glycine or proline do. Also, it is believed that it does not impose extreme electrostatic or steric effects (Cunningham and Wells, 1989; Lefevre et al., 1996). The pET28a/*AsPXG3*, the construct carrying the AsPXG3 gene which showed the highest specific activity in the *in vitro* assay was used as the template in the overlapping-PCR. Site directed mutagenesis was accomplished by using mutagenic primers and flanking primers for PCR. For each mutation, two mutagenic primers, i.e. forward and reverse mutagenic primers, were designed, as shown in Table 7, to replace 1-3 nucleotides at the target site with nucleotides that codes for alanine. DM571 and DM 581 representing T7 promoter and T7 terminator sequences, respectively, in the pET28a vector were used as two flanking primers in the overlapping PCR. The overlapping-PCR was performed under two steps and in each step 5-10 ng of template was used for amplification using 0.25 µl of NEB's Q5® high-fidelity polymerase. PCR amplification was carried out in 35 cycles. The first 10 cycles were run with an anneal temperature (T_m) 3°C below the recommended T_m of the primer pair, and the next 25 cycles were run at the anneal temperature close to optimal T_m.

The first step of overlapping PCR was carried out using the primer DM571 and the reverse mutagenic primer to amplify the 5'-part of the gene, and DM581 and the forward mutagenic primer to amplify the 3'-part of the gene, respectively. Then the mixture of the two amplicons were used as the template to generate the full length gene containing specific mutations in the second PCR with primers DM571 and DM581 (Figure 18).

Table 7: Mutagenic forward and reverse primers used in overlap PCRs. Modified codons replaced by alanine are underlined and the nucleotide changes are indicated in bold. In each primer 'F' and 'R' denotes forward and reverse orientation, respectively.

Mutant Name	Primer Name	Primer sequence (5' to 3')	Tm (^{0}C)
M1-(29L/A)	IBmF1	GACGCCG**GC**GCAGCAGCAC	66.7
	IBmR1	GTGCTGCTG**C****GC**CGGCGTC	
M2-(30Q/A)	IBmF2	GCCGCTG**GC**GCAGCACGC	67.7
	IBmR2	GCGTGCTG**C****GC**CAGCGGC	
M3-(31Q/A)	IBmF3	CTGCAG**GC**GCACGCCGCCT	68.7
	IBmR3	AGGCGGCGTG**C****GC**CTGCAG	
M4-(32H/A)	IBmF4	CTGCAGCAG**GC**CGCCGCCTTC	69.4
	IBmR4	GAAGGCGGCG**GC**CTGCTGCAG	
M5-(97I/A)	IBmF5	CATAAAAAAC**GC**CCACAAAGGCAAGC	64.1
	IBmR5	GCTTGCCTTTGTG**GGC**GTTTTTTATG	
M6-(98H/A)	IBmF6	AAACATC**GC**CAAAGGCAAGCATGG	64.6
	IBmR6	CCATGCTTGCCTTT**GGC**GATGTTT	
M7-(99K/A)	IBmF7	CATCCAC**GCA**GGCAAGCATGG	63
	IBmR7	CCATGCTTGCC**TGC**GTGGATG	
M8-(101K/A)	IBmF8	CAAAGGC**GC**GCATGGGAGTG	63.1
	IBmR8	CACTCCCATG**C****GC**GCCTTTG	
M9-(102H/A)	IBmF9	CAAAGGCAAG**GCC**GGGAGTGATAC	63.4
	IBmR9	GTATCACTCCC**GGC**CTTGCCTTTG	
M10-(103G/A)	IBmF10	GGCAAGCATG**C**GAGTGATACAGACA	62.1
	IBmR10	TGTCTGTATCACTC**G**CATGCTTGCC	
M11-(104S/A)	IBmF11	CAAGCATGGG**GCC**GATACAGACAC	63.2
	IBmR11	GTGTCTGTATC**GGC**CCCATGCTTG	
M12-(105D/A)	IBmF12	GCATGGGAGTG**CC**ACAGACACG	62.2
	IBmR12	CGTGTCTGT**GG**CACTCCCATGC	

All primers were synthesized by Sigma-Aldrich Canada Ltd. Oakville, Ontario

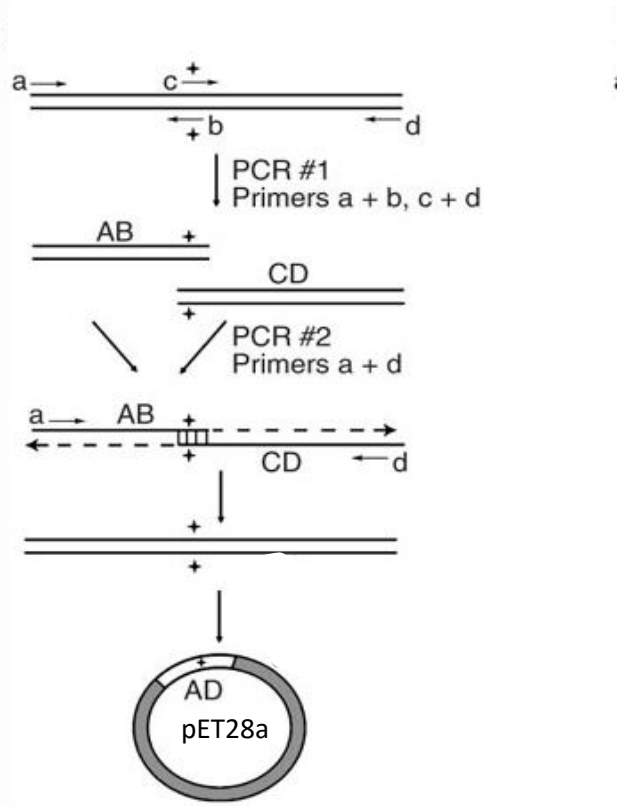


Figure 18. Site-directed mutagenesis by overlap PCR (adapted from Heckman and Peace, 2007). c and b represent the forward and reverse mutagenic primers, respectively, whereas a and d are forward and reverse flanking primers. The mutation of interest is indicated by the cross mark on the template. AB and CD represent the first and second PCR-intermediate products carrying the mutation of interest. AD is the whole gene carrying the mutation of interest.

The 1st and 2nd intermediate products separated by agarose gel electrophoresis were cut and eluted using EZ-10 spin column DNA gel extraction kit (Bio Basic Canada Inc., Mississauga, Ontario). An equal amount of gel purified intermediate products (5-10 ng) were mixed as templates for overlapping PCR to amplify the full length mutated genes (M1-M12) at the expected size of 1 Kb in length. Expected full length PCR products were then digested with *Eco*RI and cloned into the *Eco*RI site of pET28a. The recombinants carrying mutated AsPXG3 were then transformed into *E. coli* (TOP10) (Invitrogen Canada Inc., Burlington, Ontario) and selected on LB containing kanamycin (50 µg/mL) as mentioned in Study 2. Positive transformants carrying the mutated gene were screened by PCR using a forward primer, i.e. T7, and reverse gene specific primer IBW4 to check for the presence and correct orientation of the gene. The selected clone was sent to NRC (National Research Council of Canada, Saskatoon, Saskatchewan) for sequencing to confirm the mutation in the plasmid.

3.3.4.3 Functional analysis of AsPXG3-mutants by *in vitro* assays

Prior to functional analysis, the nucleotide sequences of all mutated genes in the mutant constructs were aligned with *AsPXG3*/wild type using the DNAstar program (Lasergene 9 Core Suite, Madison, USA) to confirm they carry the specific mutations. For functional analysis of the PXG in *E. coli*, the constructs were first transformed into Top 10 cells and selected transformants carrying mutations were then transformed into an expression host strain Rosetta2(DE3)pLysS using the electroporation method with an Electroporator 2510 (Eppendorf Canada Inc., Mississauga, Ontario). The mutant transformants were selected on LB containing kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL). For the enzyme expression study, Rosetta2(DE3)pLysS/AsPXG3 mutants along with Rosetta2(DE3)pLysS/AsPXG3 wild type, were grown in a LB inducing medium with 1mM IPTG at 30 ^{o}C for 4 hours. The lysates of transformants cells were used as the enzyme source in the *in vitro* assay.

As described previously in Study 2, the peroxygenase assays were carried out in a volume of 200 µL containing 0.1 mg wild type or mutated proteins, 0.5 mM cumene hydroperoxide, 0.5 mM, 18:1-9c methyl ester (ME), 100 mM Tris-HCl (pH 7.0), 20% glycerol and 0.4% emulphogene. For quantifying the products, 3 µg of 17:0-ME was used as an internal standard. After the 15 minute incubation at 45 ^{o}C, the resulting fatty acid methyl esters (FAMEs) were extracted and dried under the N_2 gas. Dried FAMEs were dissolved in 50 µL of ethyl acetate and analyzed by gas chromatography (GC). Identity of the products was confirmed by comparing GC retention time with those of standards. The specific activity of AsPXG3 mutant enzymes was quantified and then compared with respect to that of the wild type enzyme.

3.3.4.4 Substrate specificity of AsPXG3-mutants

In order to determine the substrate specificity of selected AsPXG3/mutants, a variety of free fatty acids (FFA) such as mono and poly unsaturated fatty acids ranging from C14 to C24 were provided as the substrates in the *in vitro* assays. As mentioned before the *in vitro* assays were carried out using 0.1 mg wild type or mutant proteins at 45 ^{o}C for 15 minute with shaking in a volume of 200 µL using cumene hydroperoxide as an oxidant. Fatty acid products were quantified by adding 3 µg of internal standard C17:0 FFA to the reaction mixture. After the reaction, the FFA products were methylated by 25 µL of methanol and 50 µL of diazomethane and the resulting fatty acid methyl esters were dried again under the N_2 gas. Dried FAME were

then dissolved in 50 µL of ethyl acetate and analyzed by gas chromatography (GC) analysis as before. For assaying the function of the mutants, specific activity on each substrate was analysed and compared with that of wild type AsPXG3.

3.3.5 Results

3.3.5.1 Determination of target sites of AsPXG3 for mutagenesis

Sequence alignment of three newly-identified PXGs along with oat AsPXG1 and Arabidopsis AtPXG1 and AtPXG2 (Figure 3 in Appendix) revealed that three histidine residues were highly conserved among these sequences (Figure 19). As shown in Figure 19, in addition to three conserved histidine residues H1, H2 and H3, nine amino acids surrounding these histidines are also highly conserved. Therefore, these 12 conserved amino acids were identified as primary targets for site-directed mutagenesis to study their roles in defining the peroxyganse activity. For convenience, these target sites were named as M1 (29L/A), M2 (30Q/A), M3 (31Q/A), M4 (32H/A), M5 (97I/A), M6 (98H/A), M7 (99K/A), M8 (101K/A), M9 (102H/A), M10 (103G/A), M11 (104S/A) and M12 (105D/A), respectively, in order of the amino acids from N-terminus to C-terminus of the AsPXG3 protein.

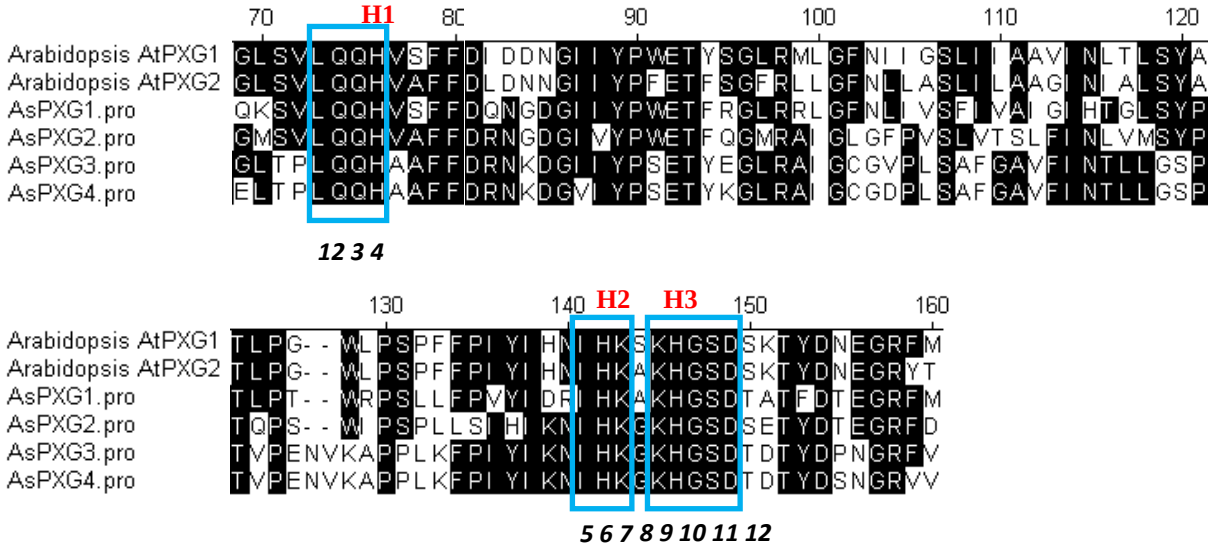


Figure 19. Partial protein sequence (69-160) alignment of *Arabidopsis* peroxygenses (AtPXG1, AtPXG2) and oat peroxygenases (AsPXG1, AsPXG2, AsPXG3 and AsPXG4). The shaded regions indicate amino acids that are identical to each other. Three highlighted areas represent the highly conserved regions surrounding three histidine-residues (H1, H2, and H3). Twelve residues in the regions chosen as the target sites for site-directed mutagenesis are labelled as 1 through 12, respectively. Broken lines in the sequence represent gaps introduced for best alignment. Each letter denotes a specific amino acid. Amino acids in conserved regions are, L; Leucine, Q; Glutamine, H; Histidine, I; Isoleucine, K; Lysine, G; Glycine, S; Serine; D; Aspartic acid.

3.3.5.2 Site-directed mutagenesis of AsPXG3

To investigate roles of these conserved amino acid residues, twelve mutant genes were generated each with a single residue being replaced by alanine. DNA fragments, constructs or strains carrying specific mutations were labelled according to the number given at the target sites. The site-directed mutagenesis was achieved through two rounds of overlapping PCR with 12 sets of mutagenic forward and reverse primers ranging 19-26 bp in length. The first round of overlapping PCR was carried out with a common forward primer DM571 and a specific mutagenic reverse primer (R1 to R12) to amplify the first intermediate fragments, and with a specific mutagenic forward primer (F1 to F12) and a common reverse primer DM581 to amplify the second intermediate fragments. The second round of the overlapping PCR was carried out using the common forward primer DM571 and the common reverse primer DM581, and a mixture of the two fragments as the template. The amplification products of the two consecutive

overlap PCRs with the expected size (Figure 20 A and B) were cloned and sequenced to confirm the individual mutations.

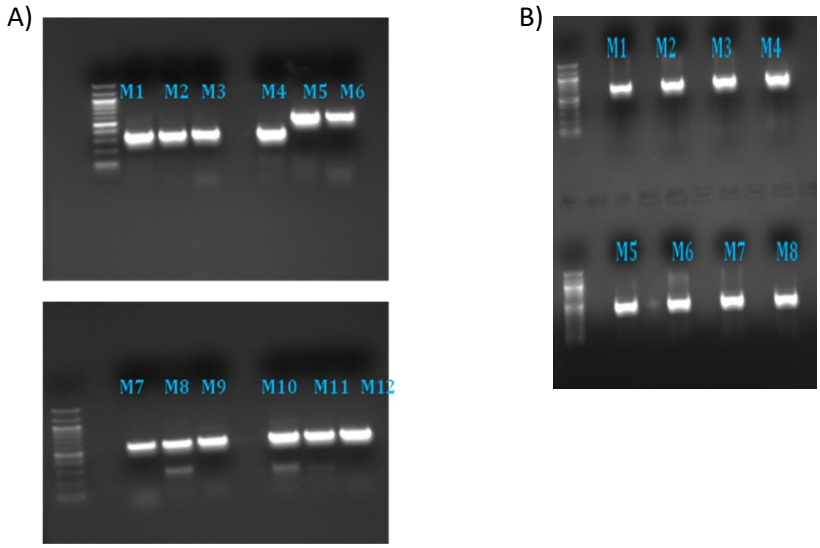


Figure 20. Overlapping PCR for site-directed mutagenesis. A) Agarose electrophoresis gel images showing the intermediate products of the 1st round PCR at the expected size of 300-350bp, 500-550bp and 550-575bp for M1-M4, M5-M7 and M8-M12, respectively. B) Agarose electrophoresis gel image showing products of the 2nd round PCR at the expected size of 650-675bp and 450-475bp for M1-M4 and M5-M8, respectively. The marker is 100bp DNA ladder.

3.3.5.3 Functional analysis of AsPXG3 mutants by *in vitro* assays

For functional analysis of these AsPXG3 mutants, twelve constructs each with a single amino acid substitution in the AsPXG3 gene were expressed in *E.coli* for *in vitro* assays. The *E.coli* transformant expressing the wild type AsPXG3 (wt) was used as the reference. Accordingly, the PXG activity of each mutant was compared to that of the wild type to determine the effect of each mutation.

In the presence of cumene hydroperoxide as an oxygen donor and oleic acid methyl ester as a substrate, the wild type AsPXG3 produced 9,10-epoxystearic acid (9,10-epoxy-18:0) methyl ester at specific activity of 47.84 µmol/mg/min. Replacement of the two conserved histidines by alanine (M4-H1, M9-H3) resulted in complete loss of the peroxygenase activity, while replacement of another conserved histidine (M6-H2) resulted in only slight reduction of the

activity by about 24%, indicating that H1 and H3, but not H2, are absolutely essential for the PXG (Figure 21 and Figure 4 in Appendix).

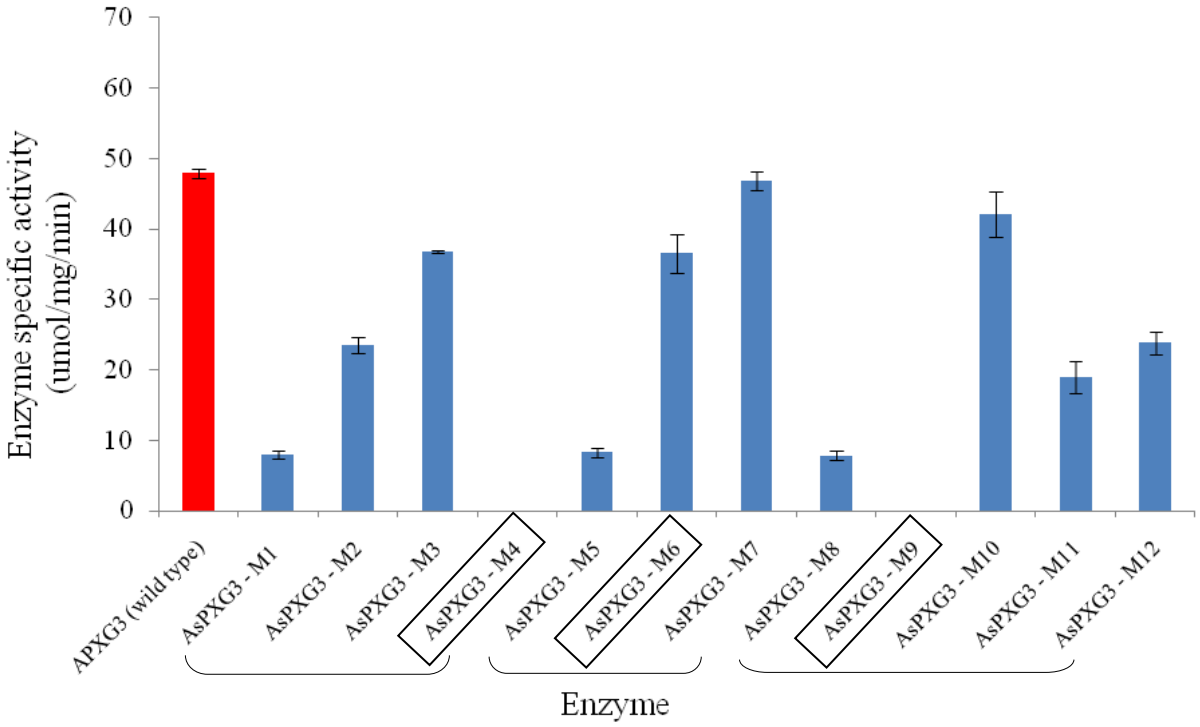


Figure 21. Comparison of specific PXG activities of twelve mutants (M1 to M12) and the wild type AsPXG3. M4 (H1/A) and M9 (H3/A) did not show any enzymatic activity in the *in vitro* assay. The assays were conducted in the presence of cumene hydroperoxide as an oxygen donor and oleic acid methyl ester as a substrate. Three brackets indicate three conserved regions according to sequence alignment. Three biological replicates were used for each treatment. Data are presented as Mean ± SD

Among the nine conserved amino acids surrounding the three histidines, replacement of glutamine at position 31 (M3), lysine at position 99 (M7) and glycine at position 103 (M10) did not considerably alter the PXG activity, with 77% to 98% of the wild type activity retained. However, substitution of leucine at position 29 (M1), isoleucine at position 97 (M5), and lysine at position 101 (M8) with alanine resulted in drastic reduction of the enzymatic activity by 83 to 84% compared to the wild type. In addition, substitution of glutamine at position 30 (M2), serine at position 104 (M11) and aspartic acid at position 105 (M12) also resulted in reduction of the activity by 50% to 60% (Table 8).

Table 8: Specific activity of the wild type AsPXG3 and mutants. Reductions in percentage of the specific activity relative to the wild type (wt) are presented. Activity not detected is denoted by '– '.

Mutants	Specific activity ± SD (µmol/mg/min)	Approximate reduction in specific activity relative to wild type (%)
AsPXG3 (wt)	47.84 ± 0.6	-
AsPXG3 - M1-(29L/A)	7.95 ± 0.5	83
AsPXG3 - M2-(30Q/A)	23.51 ± 1.0	51
AsPXG3 - M3-(31Q/A)	36.77 ± 0.1	23
AsPXG3 - M4-(32H1/A)	-	100
AsPXG3 - M5-(97I/A)	8.26 ± 0.6	83
AsPXG3 - M6-(98H2/A)	36.52 ± 2.7	24
AsPXG3 - M7-(99K/A)	46.82 ± 1.3	02
AsPXG3 - M8-(101K/A)	7.85 ± 0.6	84
AsPXG3 - M9-(102H3/A)	-	100
AsPXG3 - M10-(103G/A)	42.00 ± 3.2	12
AsPXG3 - M11-(104S/A)	18.90 ± 2.2	60
AsPXG3 - M12-(105D/A)	23.85 ± 1.6	50

Number of replicates in each treatment = 3

3.3.5.4 Substrate specificity of AsPXG3 mutants

As seen above, replacement of the two conserved histidines (H1 and H3) of AsPXG3 by alanine resulted in complete loss of the enzyme activity, indicating they are absolutely essential for PXG activity. Substitution of the leucine, isoleucine and lysine residues (M1, M5 and M8) also led to drastic reduction of the activity by more than 80%, while mutation of the other residues altered the activity less drastically. This result indicates that in addition to the two histidines, the three residues might play critical roles in defining PXG activity. Therefore, to further clarify roles of these amino acid residues in AsPXG3, possible changes in substrate specificity of the three mutants were analyzed using eight different fatty acids in the *in vitro* assays.

As shown in Figure 22, all three mutants exhibited drastic reduction of the activity towards the monounsaturated fatty acids tested. Relative to that of the wild type, the activity of the mutants was reduced by 62 to 92% on the monounsaturated fatty acids with only slight difference within each substrate. When C14:1, C16:1, C18:1, C20:1, C22:1 and C24:1 were used as the substrates in *in vitro* assays, the mutants could retain the activity at about 28-41%, 34-38%, 13-22%, 17-31%, 8-17% and 11-17% of the wild type activity, respectively (Table 9). As C22:1 was the least preferred substrate for the wild type enzyme, all three mutants also exhibited the most drastic reduction of the activity on this very long chain monounsaturated fatty acid (Table 1 in Appendix).

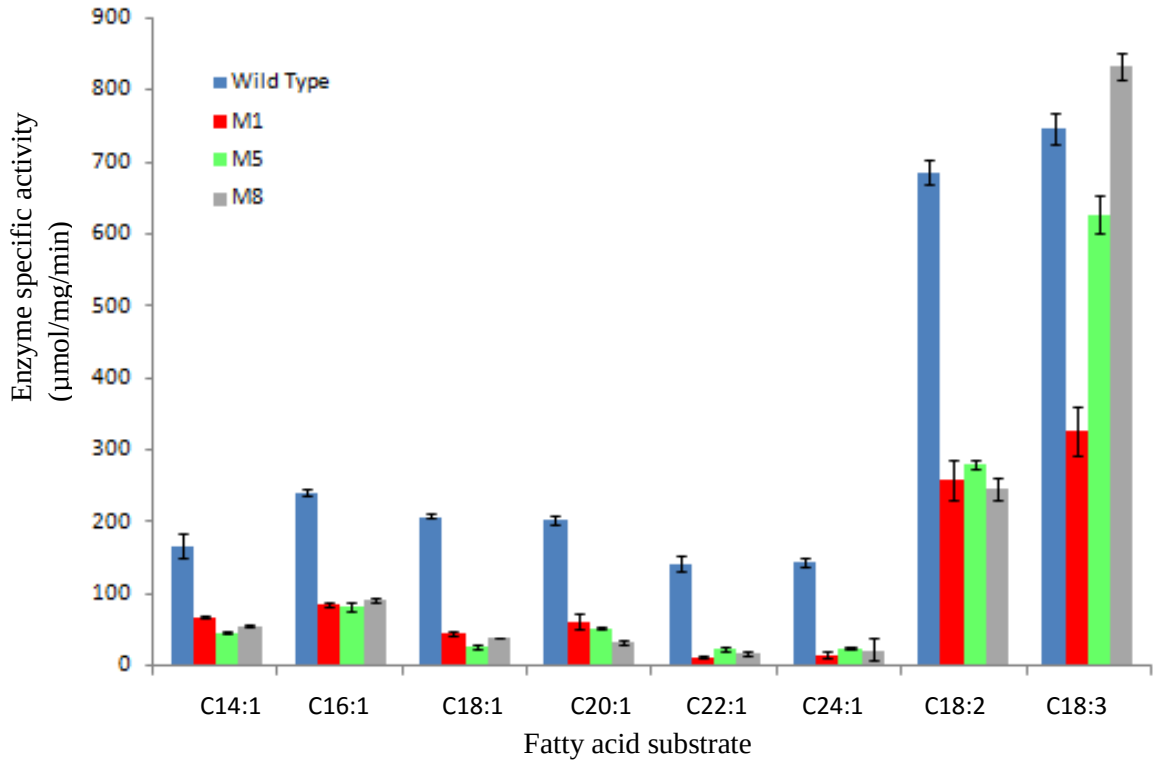


Figure 22. Specific PXG activities of three mutants of AsPXG3, M1, M5 and M8, on eight different free fatty acids. All three mutants exhibited reduced activity towards all monounsaturated fatty acids, but responded to polyunsaturated fatty acids very differently in the *in vitro* assays. The assays were conducted in the presence of cumene hydroperoxide as an oxygen donor and free fatty acids as substrates. Three biological replicates were used for each treatment. Data are presented as Mean ± SD.

Table 9. Specific activity of the three mutants M1, M5 and M8 on different fatty acids relative to the wild type.

Substrate	Wild type activity ± SD (μmol/mg/min)	Approximate activity of the mutants relative to wild type (%)		
		M1 (29L/A)	M5 (97I/A)	M8 (101K/A)
C14:1-9*c*	167 ±16.6	41	28	34
C16:1-9*c*	240 ± 4.4	35	34	38
C18:1-9*c*	208 ± 3.0	22	13	19
C20:1-11*c*	203 ± 5.3	31	26	17
C22:1-13*c*	143 ± 11.3	8	17	12
C24:1-15*c*	144 ± 6.7	11	17	16
C18:2-9*c*,12*c*	685 ± 17.2	37	41	36
C18:3-9*c*,12*c*,15*c*	746 ± 22.0	44	84	112

Number of replicates in each treatment = 3

Interestingly, the three mutants displayed the activity on polyunsaturated fatty acids very differently. Like the wild type, M1, M5 and M8 mutants had higher catalytic activity on linoleic acid and linolenic acid over monounsaturated fatty acids (Figure 21). On linoleic acid, the three mutants exhibited similar reduction of the activity by about 60% (Figure 21 and Table 7). On linolenic acid, however, the mutants behaved very differently. Replacement of leucine by alanine (M1) resulted in 56% reduction of the activity, replacement of isoleucine (M5) reduced the activity by only 16%, whereas lysine replacement (M8) resulted in even an increase of the PXG activity by 12% (Figure 5 and 6 in Appendix).

3.3.6 Discussion

From the results shown above, it could be concluded that among three conserved histidine residues in AsPXG3, H1 and H3 are absolutely essential for the PXG activity, as substitution of the two residues results in complete loss of the activity. Besides the two histidines, leucine at position 29, isoleucine at position 97, and lysine at position 101 would also play very important roles in defining the activity, as substitution of these amino acids reduces the enzymatic activity by more than 83%. Glutamine at position 30, serine at position 104 and

aspartic acid at position 105 would also have some roles for the enzymatic activity, as substitution of these amino acids resulted in reduction of the original activity by approximately half. On the other hand, glutamine at position 31, lysine at position 99 and glycine at position 103 might not have significant roles, as replacement of these amino acids results in only slight reduction of the enzyme activity (less than 25%).

Hemoproteins are a group of proteins carrying heme as a prosthetic group. They are abundant in biological systems and exhibit diverse biological functions such as gas transportation/storage and electron transmission, which are built on interaction with the prosthetic heme group (Schneider et al., 2007; Li et al., 2010). According to structural analysis of heme proteins, five different amino acids (histidine, methionine, cysteine, tyrosine and lysine) can serve as axial ligands to the heme iron, of which histidine is the most prominent. For instance, in cytochrome P450, a cysteine thiolate was identified to be the heme axial ligand, whereas in a horseradish peroxidase, a histidine is coordinated to the heme. In peroxidase, heme iron is pentacoordinated with histidine-170 being the axial ligand and histidine-42 being the catalytically relevant residue in the distal heme side (Gajhedi et al., 1997; Hanano et al., 2006).

Plant peroxygenase belongs to a hemoprotein where histidine residues are critical for the enzyme activity. It catalyzes epoxidation of unsaturated fatty acids by transferring an oxygen atom from an oxo-feryl intermediate in the prosthetic group to a substrate (Hanano et al., 2006). The oxo-feryl intermediate plays an important role in the catalytic mechanism; however, formation of the oxo-feryl intermediate is dependent on the architecture of the heme binding site or the heme binding ligands in a PXG. Site-directed mutagenesis on AtPXG1 revealed two conserved histidine residues, His-70 and His-138, as the most critical residues for peroxygenase activity, of which His-70 was proposed as a heme axial ligand (penta-coordinated heme), as involvement of two histidines in the heme coordination (hexa-coordinated heme) in plant PXGs has not been confirmed (Hanano et al., 2006). Consistently, site-directed mutagenesis in the present study also showed two histidine residues are absolutely essential for the oat PXG activity. However, whether these two histidines are all involved in heme-binding remains to be clarified.

Heme proteins possess a pocket that strongly binds a heme group. Thus, availability of heme-binding ligands and the structural feature for heme binding in the pocket are important for a heme-containing enzyme. Heme is a cofactor consisting of an Fe^{2+} ion in the centre of a large heterocyclic organic ring called a porphyrin, which is made up of four pyrrolic groups joined together by methine bridges (Figure 23).

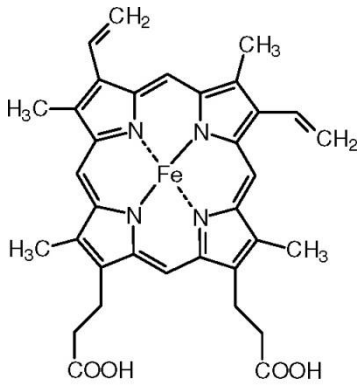


Figure 23. Schematic diagram of heme group with Fe^{2+} ion in the centre of a large heterocyclic organic ring called a porphyrin. The iron ion, which is the site of oxygen binding, coordinates with the four nitrogens in the center of the ring. The ring structure and iron lie in one plane. Source www.bio.davidson.edu

The heme-binding pocket is generally enriched in aromatic and non-polar amino acids. In this context, conserved hydrophobic residues surrounding the histidines in AsPXG3 would play important roles in maintaining a hydrophobic environment in the pocket interacting with the heme-ring structure or substrates (Smith et al., 2010; Li et al., 2011). Consistent with the hypothesis, replacement of highly hydrophobic residues leucine at position 29 and isoleucine at position 97 (hydrophobic index at pH 7= 97 and 99 respectively) with less hydrophobic alanine (hydrophobic index at pH 7= 41) might have altered the hydrophobic interactions with the heme ring. Further, it might have adversely affected the hydrophobic interactions between the substrate and the enzyme, together resulting in drastic reduction of the enzyme activity. However, in this regard, different effect of electrically charged residue lysine replacement on the enzyme activity toward different substrates is not easily explainable.

Substrate specificity assays showed that like the other mutants, lysine replacement (M8) results in drastic decrease of the activity on monounsaturated fatty acids, but unlike the other

mutants, the substitution leads to even an increase of PXG activity on linolenic acid, implying this residue might be involved in interaction with double bonds of the substrate. Alternatively, this specific mutation might have increased dynamics of oxo-feryl intermediate for an efficient oxygen transmission to the substrate.

Previous structural analysis revealed that the size and shape of heme-binding pocket formed by the surrounding amino acids are important for fixing a heme at the proper position (Li et al., 2011). The proteins where the heme group plays an electron transfer role have a small hydrophobic pocket and its shape is closely complementary to the shape of the heme group. However, in proteins like plant PXG with co-oxidation capacity where the heme group is part of the active site, the heme-binding pocket tends to be larger with a variety of shapes (Smith et al., 2010). Alteration of glutamine at position 30, serine at position 104 and aspartic acid at position 105 results in reduction of the activity by 50% to 60%, implying these amino acids might be involved, to some extent, in maintaining the shape and size of the heme binding pocket. After alanine substitution, the structure in these mutants may have adopted a different conformation from the wild type to fill the void formed by removal of a bulky side chain with other residues (Paoli et al., 2002). Therefore, replacement of these amino acids with relatively smaller residues would lead to considerable reduction of the enzyme activity as a result of the physical structure change in the catalytic pocket.

4. GENERAL DISCUSSION AND CONCLUSIONS

The demand for oat as a food and feed has gradually increased in past years, due to its particular chemical composition and recognized nutritive value. Oat is characterized by good taste, dietetic property and metabolic activity in the body. In comparison with other cereals, oat grain is unusual, as it contains a larger amount of protein, oil and β-glucan. The characteristic feature is its good amino acid composition of the proteins with high nutritive value and a good source of essential unsaturated fatty acids with a high level of fat. In addition, there are minor components such as avenanthramides, a class of polyphenolics unique to oat, with high antioxidant properties. All these characteristics make oat a highly nutritive grain for human and animals (Peterson, 2001; Peltonen-Sainio et al., 2004).

Besides the excellent nutritional value, oat has also been known to produce several epoxy fatty acids synthesized through a peroxygenase pathway (Hamberg and Hamberg, 1996). Interestingly, this class of unusual fatty acids in oat has some potential industrial applications. However, due to their oxygenated properties, oat breeding and oat food industry are seeking an effective strategy that can minimize the formation and/or accumulation of the epoxy fatty acids in oat or oat-based food products. In this regard, the importance of studying oat peroxygenase is apparent in enhancing the quality of oat for novel food applications. For instance, extensive understanding about oat peroxygenases involved in biosynthesis of epoxy fatty acids can help formulate an effective strategy to manipulate epoxy fatty acid content in oat.

This thesis research started with identification of additional cDNAs encoding fatty acid peroxygenase from oat by searching oat EST databases prepared from leaves, roots, and developing seeds at different stages that are publically available from the NCBI using previously identified AsPXG1 as the query sequence. This resulted in identifying five additional consensus contig sequences. The study was then focused on two contig sequence *AsPXG2* and *AsPXG3* by cloning and characterization of their full length ORFs. Retrieval of the missing ends of *AsPXG2* and *AsPXG3* was conducted using cDNA synthesized from the total RNA of developing seeds of CDC Dancer. Sequencing of amplified products confirmed that three *AsPXGs*, *AsPXG2*, *AsPXG3* and *AsPXG4*, were obtained (retrieving *AsPXG3* resulted in two cDNA sequences).

Sequence analysis showed a calcium-binding motif, a transmembrane domain and three histidine residues were conserved among these putative oat PXGs. According to topology prediction and phylogenetic analysis, AsPXG2 is a Class I caleosin, whereas AsPXG3 and AsPXG4 belong to Class II caleosins.

The second study focused on functional characterization of the candidate PXG genes. To do so, all three sequences were sub-cloned into a bacterial/yeast expression vector and the recombinant plasmids were introduced to Rosetta2(DE3)pLysS/*Pichia pastoris.* The lysates of transformants were used as enzyme sources for *in vitro* assays and function of the genes was determined by gas chromatography analysis of the assay products. The analysis showed that all three putative PXGs code for functional peroxygenase, and in the presence of cumene hydroperoxide, these PXGs can epoxidize oleic acid methyl ester, but AsPXG3 had the highest peroxygenase activity. Then, substrate specificity of oat AsPXG3 was determined using a variety of free unsaturated fatty acids as substrates in *in vitro* assays. Consistent with previous findings (Blee et al., 2012), polyunsaturated fatty acids proved to be the preferred substrates. This further confirmed that AsPXG3is a Class II caleosin newly identified in oats.

Finally, structure-function relationship of an oat PXG gene was studied by introducing single point mutations into AsPXG3 through overlapping PCR approach. Twelve mutants of AsPXG3 were generated by substituting twelve conserved residues, three histidines and nine residues surrounding these histidines in AsPXG3, with alanine individually. The *in vitro* enzymatic assays of these mutants were used to identify crucial residues of the PXG enzyme for the enzymatic properties. The mutagenesis analysis showed that among twelve conserved residues, two histidines: H1 and H3, are absolutely essential for the PXG activity. They might be involved in heme binding, which is critical for plant PXG activity (Hanano et al., 2006; Blee et al., 2012). Besides two histidines, leucine (29), isoleucine (97) and lysine (101) were also very important, glutamine (30), serine (104) and aspartic (105) were important for the PXG activity, while glutamine (31), lysine (99) and glycine (103) prove to be insignificant for the catalytic activity. It is believed that, apart from two histidines, the important residues might be involved in shaping the stereospecific structure of the heme-binding pocket and/or maintaining specific interactions of the heme-ring with substrates.

Molecular genetics and genomics can substantially contribute to improvement of oat and oat quality in several ways. For instance, they increase our knowledge on the genetic background

of oat quality, which can lead to effective breeding strategies for quality improvement (Valentine et al., 2011). Using such strategies, plant breeders have the opportunity to alter chemical composition of oat grain to meet new nutritional demands. Overall, this thesis has provided new pertinent information about epoxy fatty acid biosynthesis and catalytic activity of peroxygenase in oat. The data from the research can be used to develop functional molecular markers for oat breeding to manipulate the PXG activity and epoxy fatty acid content.

5. REFERENCES

Agriculture and Agri-Food Canada, 2010. Market outlook report.Accessed online 2012 May20< http://www.agr.gc.ca/pol/mad-dam/index_e.php?s1=pubs& s2=rmar&s3=php&page=rmar _02_03_2010-08-03>

Agriculture & Agri-Food Canada (2012) Canadian Agricultural Biodiversity. National Collections. Accessed online 2014 March 29<http://www4.agr.gc.ca/AAFC-AAC/display-afficher.do?id=1273845048950&lang=eng>

Bafor, M., Smith, M. A., Jonsson, L., Stobart, K. and Stymne, S. (1993). Biosynthesis of vernoleate (*cis*-12-epoxyoctadeca*cis*- 9-enoate) in microsomal preparations from developing endosperm of *Euphorbia lagascae*. *Archives of Biochemistry and Biophysics.* 303:145–151.

Banas, A., Debski, H., Banas, W., Heneen, W.K., Dahlqyist, A., Bafor, M., Gummeson, P.O., Marttila, S., Ekman, Å., Carlsson, A.S. and Stymne, S. (2007). Lipids in grain tissues of oat (*Avena sativa*): differences in content, time of deposition and fatty acid composition. *Journal of Experimental Botany.* 58:2463-2470.

Blee, E. and Durst, F. (1987). Hydroperoxide-dependent sulfoxidation catalyzed by soybean microsomes. *Archives of Biochemistry and Biophysics*. 254:43–52.

Blee, E., Wilcox, A. L., Marnett, L. J. and Schuber, F. (1993). Mechanism of reaction of fatty acid hydroperoxides with soybean peroxygenase. *Journal of Biological Chemistry*. 268: 1708–1715.

Blee, E., Flenet, M., Boachon, B. and Fauconnier, M. L. (2012). A non-canonical caleosin from Arabidopsis efficiently epoxidizes physiological unsaturated fatty acids with complete stereoselectivity. *Federation of European Biochemical Societies Journal.* 279:3981- 3995.

Brennan, C. S. and Cleary, L. J. (2005). Review: The potential use of cereal (1→3,1→4)-β-D-glucans as functional food ingredients. *Journal of Cereal Science*. 42:1-13.

Bradford, M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochemistry*. 72:248–254.

Buckeridge, M. S., Rayon, C., Urbanowicz, B., Tine, M. A. S. and Carpita, N. C. (2004). Mixed linkage (1/3), (1/4)-β-D-glucans of grasses. *Cereal Chemistry*. 81:115–127.

Budziszewski, G. J., Croft, K. P. C. and Hildebrand, D. F (1996). Uses of biotechnology in modifying plant lipids. *Lipids.* 31:557-569.

Butt, M. S., Tahir-Nadeem, M., Khan, M. K. I., Shabir, R. and Butt, M. S. (2008). Oat: unique among the cereal. *European Journal of Nutrition*. 47:68–79.

Cahoon, E. B., Ripp, K. G., Hall, S. E. and McGonigle, B. (2002). Transgenic production of epoxy fatty acids by expression of a cytochrome P450 enzyme from *Euphorbia lagascae* seed. *Plant Physiology*. 128:615–624

Carlsson, A. S., Clayton, D., Salentijn, E., Toonen, M., Stymne, S., Dyer, W. J. and Bowles, D. (2007). In: *Oil crop platforms for industrial uses*, CNAP, University of York: Berks, United Kingdom, p:155.

Chehab, E. W., Perera, J. V., Gopalan, B., Theg, S. and Dehesh, K. (2007). Oxylipin pathway in rice and *Arabidopsis*. *Journal of Integrative Plant Biology*. 49:43–51.

Coffman, F. A (1977). The species of *Avena* Technical bulletin no. 1516: Oat history, identification, and classification. Agricultural Research Service, United States Department of Agriculture. Washington, D.C.

Cunningham, B. C. and Wells, J. A. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*. 244:1081–1085.

Delaney, B., Nicolosi, R. J., Wilson, T. A., Carlson, T., Frazer, S., Zheng, G. H., Hess, R., Ostergren, K., Haworth, J. and Knutson, N. (2003). Beta-glucan fractions from barley and oats are similarly antiatherogenic in hypercholesterolemic Syrian golden hamsters. *Journal of Nutrition.* 133: 468-75.

Deshpande, S. S. and Damodaran, S. (1990). Chapter 3: Food legumes: chemistry and technology. In: *Advances in Cereal Science and Technology*, Volume X, Pomeranz, Y. American Association of Cereal Chemists, St. Paul, Minnesota. pp 147-241.

Doehlert, D. C., Angelikousis, S. and Vick, B. (2010). Accumulation of oxygenated fatty acids in oat lipids during storage. *Cereal Chemistry.* 87:532-537.

Dykes, L. and Rooney, L. W. (2007). Phenolic compounds in cereal grains and their health benefits. *Cereal Foods World.* 52:105-111.

Ekstrand, B., Gangby, I. and Akesson, G. (1992).Lipase activity inoats-distribution, pH-dependence, and heat inactivation. *Cereal Chemistry*. 69:379-381.

Erkkila, A. T., Herrington, D. M., Mozaffarian, D. and Lichtenstein, A. H. (2005). Cereal fiber and whole-grain intake are associated with reduced progression of coronary-artery atherosclerosis in postmenopausal women with coronary artery disease. *Amerrican Heart Journal.* 150:94-101.

Fuchs, C. and Schwab, W. (2013). Epoxidation, hydroxylation and aromatization is catalyzed by aperoxygenase from *Solanum lycopersicum*. *Journal of Molecular Catalysis B: Enzymatic*. 96: 52–60.

Fulcher, R.G. (1986). Morphological and chemical organization of the oat kernel. In: *Oats: Chemistry and Technology*. Webster, F. H. American Association of Cereal Chemists, St Paul, Minnesota, pp. 47–74.

Gajhede, M., Schuller, D. J., Henriksen, A., Smith, A. and Poulos, T. (1997). Crystal structure of horseradish peroxidase C at 2.15 Å resolution. *Nature Structural and Moecular Biology.* 4: 1032–1034.

Hamberg, M. and Hamberg, G. (1990). Hyrdroperoxide dependent epoxidation of unsaturated fatty acids in the broad bean (*Viciafaba,* L.). *Archives of Biochemistry and Biophysics.* 283:409-416.

Hamberg, M. and Hamberg, G. (1996). Peroxygenase-catalyzed fatty acid epoxidation in cereal seeds (sequential oxidation of linoleic acid into 9(S),12(S),13(S)-trihydroxy-10(E)-octadecenoic acid). *Plant Physiology.* 110:807–815.

Hanano, A., Burcklen, M., Flenet, M., Ivancich, A., Louwagie, M., Garin, J. and Blee, E. (2006). Plant seed peroxygenase is an original heme-oxygenase with an EF-hand calcium binding motif. *Journal of Biological Chemistry.* 281:33140–33151.

Health Canada (2012). Fats: The Good the Bad and the Ugly. Healthy Living. Accessed online 2014 January 02 <http://www.hc-sc.gc.ca/hl-vs/iyh-vsv/med/fats-gras-eng.php>

Hemoglobin. Accessed online 2015 May 15 <http://www.bio.davidson.edu/Courses/Molbio/MolStudents/spring2010/Hua/Hemoglobin.htm>l

Heneen, W. K., Karlsson, G., Brismar, K., Gummeson, P. O., Marttila, S., Leonova, S., Carlsson, A. S., Bafor, M., Banas, A., Mattsson, B., Debski, H. and Stymne, S. (2008). Fusion of oil bodies in endosperm of oat grains. *Planta*. 228:589-599.

Heneen, W. K., Banas, A., Leonov, S., Carlsson, A. S., Marttila, S., Debski, H. and Stymne, S. (2009). The distribution of oil in the oat grain. *Plant Signaling and Behavior*. 4:55-56.

Heydanek, M. G. and McGorrin, R. J. (1981). Gas chromatography-mass spectroscopy investigations on the flavor chemistry of oat groats. *Journal of Agricultural and Food Chemistry.* 29:950–954.

Heydanek, M. G. and McGorrin, R. J. (1986). Oat flavour chemistry: principles and prospects. In: *Oats: Chemistry and Technology*. Webster, F. H. American Association of Cereal Chemists, St Paul, Minnesota, pp. 335–369.

Holland, J. B., Frey, K. J. and Hammond, E. G. (2001). Correlated responses of fatty acid composition, grain quality and agronomic traits to nine cycles of recurrent selection for increased oil content in oat. *Euphytica*. 122:69-79.

Invitrogen: *Pfx50*™ DNA Polymerase(2006). Accessed online 2012 May 2< http://tools.invitrogen.com/content/sfs/manuals/pfx50_man.pdf>

Invitrogen: pPICZ A, B, and C Pichia expression vectors for selection on Zeocin™ and purification of recombinant proteins (2010). Accessed online 2012 May 15< http://tools.invitrogen.com/content/sfs/manuals/ppicz_man.pdf>

Ishimaru, A. and Yamazaki, I. (1977). Hydroperoxide-dependent hydroxylation involving "H_2O_2-reducible hemoprotein" in microsomes of pea seeds: a new type enzyme acting on hydroperoxide and a physiological role of seed lipoxygenase. *Journal of Biological Chemistry*. 252:6118–6124.

Kapicac, (2001). Oats – nature's functional food. *Nutrition Today*. 36:56–60.

Kurtz, E. S. and Wallo, W. (2007). Colloidal oatmeal: history, chemistry and clinical properties. *Journal of Drugs in Dermatology*. 6:167–170.

Lee, M., Lenman, M., Banaś, A., Bafor, M., Singh, S., Schweizer, S., Nilsson,R., Liljenberg, C., Dahlqvist, A., Gummeson, P., Sjödahl, S., Green, A. and Stymne, S. (1998). Identification of non-heme diiron proteins that catalyze triple bond and epoxy group formation. *Science.* 280: 915–918.

Lefevre˙ L., Remy, M. H. and Masson, J. M. (1996). Alanine-stretch scanning mutagenesis: a simple and efficient method to probe protein structure and function. *Nucleic Acid Research.* 25: 447-448.

Lehtinen, P., KiiliaÈinen, K., Lehtoma eki, L. and Laakso, S. (2003). Effect of heat treatment on lipid stability in processed oats. *Journal of Cereal Science.* 37:215-221.

Lehtinen, P. and Kaukovirta-Norja, A. (2011). Chapter 9: Oat lipids, enzymes, and quality. In: *Oats chemistry and technology, second edition*. Webster, F. H. and Wood, P. J. American Association of Cereal Chemists, St. Paul, Minnesota. pp. 143-156.

Leonova, S., Shelenga, T., Hamberg, M., Konarev, A. V., Loskutov, I. and Carlsson, A. S. (2007). Analysis of oil composition in cultivars and wild species of oat (Avena sp.). *Journal of Agricultural and Food Chemistry.* 56:7983-7991.

Lequeu, J., Fauconnier, M. L., Chammai, A., Bronner, R. and Blee, E. (2003). Formation of plant cuticle: evidence for the occurrence of the peroxygenase pathway. *Plant Journal.* 36:155-164.

Li, C. D. Rossnagel, B. G. and Scoles, G. J. (2000). Tracing the phylogeny of haploid oat *Avena sativa* with satellite DNAs. *Crop Science*. 40:1755-1763.

Li, R., Yu, K. and Hildebrand, D. F. (2010). DGAT1, DGAT2 and PDAT expression in seeds and other tissues of epoxy and hydroxy fatty acid accumulating plants. *Lipids*. 45:145–157.

Lichtenstein, A. H., Kennedy, E., Barrier, P., Danford, D., Ernst, N. D., Grundy, S. M., Leveille, G. A., Van Horn, L., Williams, C. L. and Booth, S. L. (1998). Dietary fat consumption and health. *Nutrition Review*. 56:3-19.

Liu, L., Hammond, E. G. and Nikolau, B. J. (1998). *In vivo* studies of the biosynthesis of vernolic acid in the seed of *Vernonia galamensis*. *Lipids*. 33:1217–122.

Liukkonen, K., Montfoort, A. and Laakso, S. V. (1992). Water induced lipid changes in oat processing. *Journal of Agricultural and Food Chemistry*. 40:126-130.

Lockhart, H. B. and Hurt, H. D. (1986). Nutrition of oats. In: *Oats: Chemistry and technology*, Webster, F. H. American Association of Cereal Chemists, St. Paul, Minnesota. pp. 297-308.

Loskutov, I.G. (2001). Inter-specific crosses in the genus *Avena* L. *Russian Journal of Genetics.* 37:467-475.

Marlet, J.A. (1993). Comparisons of dietary fiber and selected nutrient compositions of oat and other grain fractions. In: *Oat Bran*. Wood, P. J. American Association of Cereal Chemists St. Paul, Minnesota. pp. 49-82.

Marmesat, S., Velasco, J. and Dobarganes, M. C. (2008). Quantitative determination of epoxy acids, keto acids and hydroxy acids formed in fats and oils at frying temperatures. *Journal of Chromatographgy A.* 1211:129–134.

Matsunga, I. and Shiro, Y. (2004). Peroxide-utilizing biocatalysts: structural and functional diversity of heme-containing enzymes. *Current Opinion in Chemical Biology.* 8:127-132.

Mayfield, J. A., Fiebig, A., Johnstone, S. E., and Preuss, D. (2001). Gene families from the *Arabidopsis thaliana* pollen coat proteome. *Science.* 292:2482–2485.

Meesapyodsuk, D. and Qiu, X. (2011). A peroxygenase pathway involved in the biosynthesis of epoxy fatty acids in oat. *Plant Physiology.* 157:454-63.

Meydani, M. (2009). Potential health benefi ts of avenanthramides of oats. *Nutrition Reviews*. 67:731–735.

Molteberg, E. I., Magnus, E. M., Bjorge, J. M. and Nilsson, A. (1996). Sensory and Chemical studies of lipid oxidation in raw and heated oat flours. *Cereal Chemistry*. 73:579-587.

Murphy, D. J. (2001). The biogenesis and functions of lipid bodies in animals, plants and microorganisms. *Progress in Lipid Research.* 40:325–438.

Murphy, D. J. (2005). Lipid-associated proteins, in: Plant Lipids: Biology, Utilisation and Manipulation, Blackwell, Oxford, pp. 226–269

Naested, H., Frandsen, G. I., Jauh, G. Y., Hernandez-Pinzon, I., Nielsen, H. B., Murphy, D. J., Rogers, J. C. and Mundy, J. (2000). Caleosins: Ca^{2+}-binding proteins associated with lipid bodies. *Plant Molecular Biology*. 44:463–476.

Napier, J.A. (2007). The production of unusual fatty acids in transgenic plants. *Annual Review of Plant Biology.* 58:295-315.

Novagen pET system manual (2010). 10th Edition. Accessed online 2013 September 17< www.novagen.com>

Othman, R. A., Moghadasian, H. M. and Jones, P. T. (2011). Cholesterol-lowering effects of oat beta-glucan. *Nutrition Reviews.* 9:299-309

Paoli, M., Marles-Wright, J. and Smith, A. (2002). Structure-function relationships in heme-proteins. *DNA Cell Biology.* 21:271-280.

Partridge, M. and Murphy, D. J. (2009). Roles of a membrane-bound caleosin and putative peroxygenase in biotic and abiotic stress responses in Arabidopsis. *Plant Physiology and Biochemistry*. 47:796–806.

Peltonen-sainio, P., Kontturi, M., Rajala, A. and Kirkkari, A. M. (2004). Impact dehulling oat grain to improve quality of on-farm produced feed – 1. Hullability and associated changes in nutritive value and energy content. *Agricultural and Food Science*. 13:18–28.

Perdue, R. E., Carlson, K. D. and Gilbert, M. G. (1986). *Vemonia galamensis*, potential new crop source of epoxy acid. *Economic Botany*. 40:54-68.

Peterson, D. M. and Wood, D. F. (1997). Composition and structure of high oil oat. *Journal of Cereal Science.* 26:121-128.

Peterson, D. M. (2001). Oat antioxidants. *Journal of Cereal Science*. 33:115–129.

Piazza, G. J., Foglia, T. A. and Nunez, A. (2001). Optimizing reaction parameters for the enzymatic synthesis of epoxidized oleic acid with oat seed peroxygenase. *Journal of the American Oil Chemists Society.* 78:589-592.

Piazza, G. J., Nuñez, A. and Foglia, T. A. (2003). Isolation of unsaturated diols after oxidation of conjugated linoleic acid with peroxygenase. *Lipids*. 38:255–261.

Rajhathy, T. and Thomas, H. (1974). Cytogenetics of oats (*Avena* L.). *Miscellaneous Publications of the Genetics Society of Canada – No. 2.* The Genetics Society of Canada. Ottawa, ON.

Reynolds, S. G. (2004). Chapter 1: Fodder Oats: A world overview. In: *Fodder oats: a world overview*. Suttie, J. M. and Reynolds, S. G. *Plant Production and Protection Series No. 33.* Food and agriculture organization of the United Nations. pp. 37-43.

Rezzonico, E., Moire, L., Delessert, S. and Poirier, Y. (2004). Level of accumulation of epoxy fatty acid in *Arabidopsis thaliana* expressing a linoleic acid *Δ*12-epoxygenase is influenced by the availability of the substrate linoleic acid. *Theoretical and Applied Genetics.* 109:1077–1082.

Ruxton, C. H. S., Reed, S. C., Simpson, M. J. A. and Millington, K. J. (2007). The health benefits of omega-3 polyunsaturated fatty acids: A review of the evidence. *Journal of Human Nutrition and Dietetics.* 20:275–285.

Schrickel, D. J. (1986).Chapter 1: Oats production, value, and use. In: *Oats: Chemistry and Technology*, Webster, F. H. American Association of Cereal Chemists, Inc. St. Paul, Minnesota. pp. 1-11.

Schneider, S., Marles-Wright, J., Sharp, K. H. and Paoli, M. (2007**).** Diversity and conservation of interactions for binding heme in b-type heme proteins**.** *Natural Product Reports.* 24:621-630.

SeqMan tutorials (Mac) (2005). Accessed online 2012 May 21 <www.science.smith.edu/cmbs/SeqMan.doc>

Seqbuilder (2005). Accessed online 2012 May 25< http://www.dur.ac.uk/stat.web/Bioinformatics/ SeqBuilder.htm [May 25 2012]

Stevens, E. J., Armstrong, K. W., Bezar, H. J., Griffin, W.B. and Hampton, J. G. (2004). Chapter II – Fodder Oats: An overview. In: *Fodder oats: a world overview.* Suttie, J.M. and Reynolds, S. G. *Plant Production and Protection Series No. 33.* Food and Agriculture Organization of the United Nations.

Singh, S., Thomaeus, S., Lee, M., Stymne, S. and Green, A. (2001). Transgenic expression of a delta 12-epoxygenase gene in Arabidopsis seeds inhibits accumulation of linoleic acid. *Planta.* 212:872-879.

Smith, L. J., Kahraman, A. and Thornton, J. M. (2010). Heme proteins diversity in structural characteristics, function, and folding. *Proteins*. 78:2349-2368.

Statistics Canada. 2009. Field crop reporting areas. Accessed online 2012 May 26< http://www.statcan. gc.ca/pub/22-002-x/2009008/t032-eng.htm>

Statistic canda. 2012. Accessed online 2012 May 26< http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/prim11a-eng.htm>

Sugawara, T. and Miyazawa, T. (1999). Separation and determination of glycolipids from edible plant sources by high-performance liquid chromatography and evaporative-light scattering detection. *Lipids.* 34:1231-1237.

Thomaeus, S., Anders, S. C. and Stymne, S. (2001). Distribution of fatty acids in polar and neutral lipids during seed development in *Arabidopsis thaliana* genetically engineered to produce acetylenic, epoxy and hydroxy fatty acids. *Plant Science*. 161:997–1003.

Valentine, J., Cowan, A. A. and Marshall, A. H. (2011). Chapter 2: Oat breeding. In: *Oats chemistry and Technology*. 2nd edition. Webster, F. H. and Wood, J. American Association of Cereal Chemists St. Paul, Minnesota pp: 11-30.

Youngs, V . L. and Forsberg, R. A. (1987). Oat. In: *Nutritional Quality of Cereal Grains.* Frey, O. A. K. J. Madison, WI: American Society of Agronomy. pp. 457-499.

Zhou, M., Rbards, K., Glennie-holmes, M. and Helliwell, S. (1999). Oat lipids. *Journal of the American Oil Chemists' Society*. 76:159–169.

Zingg, J. M. and Azzi, A. A. (2004). Non-antioxidant activities of vitamin E. *Current Medicinal Chemistry.* 11:1113–1133.

Zwer, P. (2010). Chapter 1: Oats: characteristics and quality Requirements. *In: Cereal grains. Assessing and managing quality*. Wrigley, C. and Batey, I. Woodhead Publishing Series in Food Science, Technology and Nutrition. pp. 163-182.

6. APPENDICES

Figure 1: Alignment showing the amino acid comparison of *AsPXG3* and *AsPXG4*. The shaded regions indicate amino acids that are identical to each other. The non-shaded regions represent the different amino acids in the two proteins.

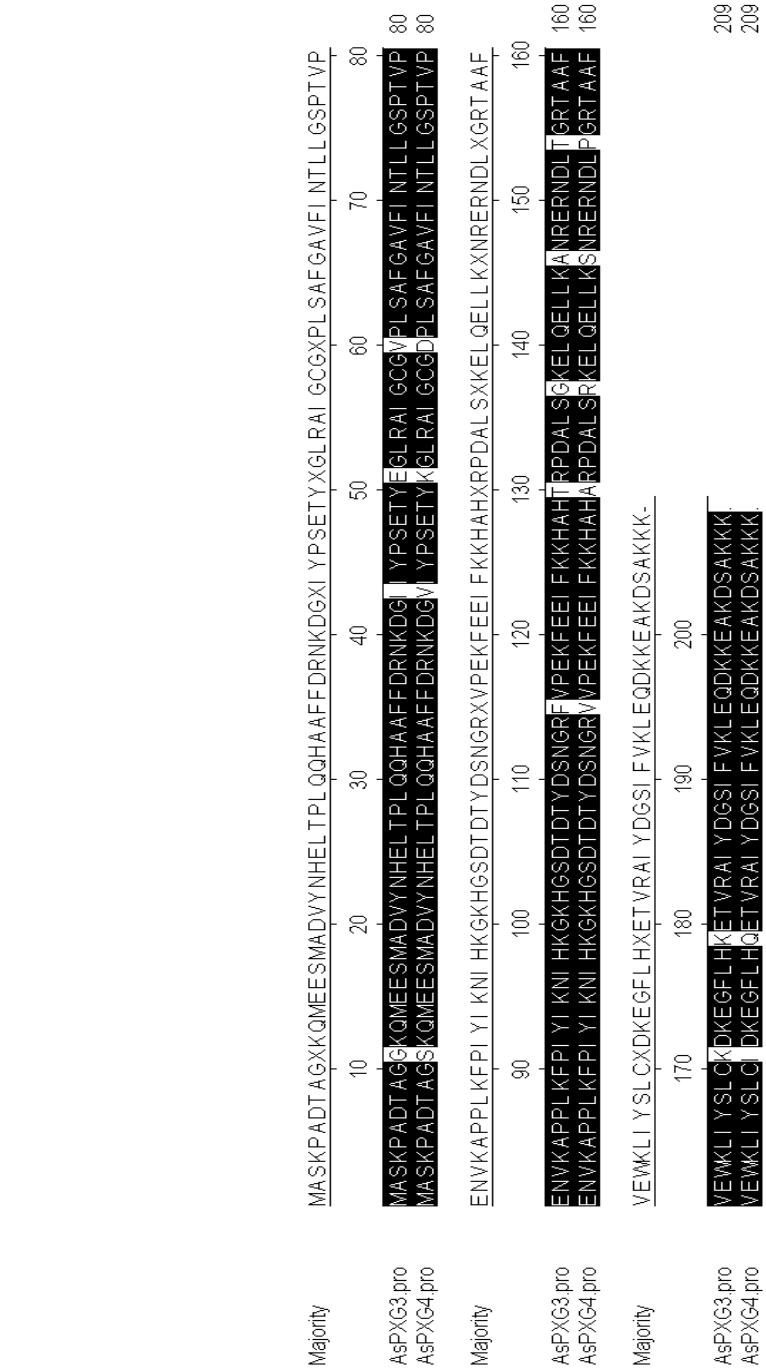
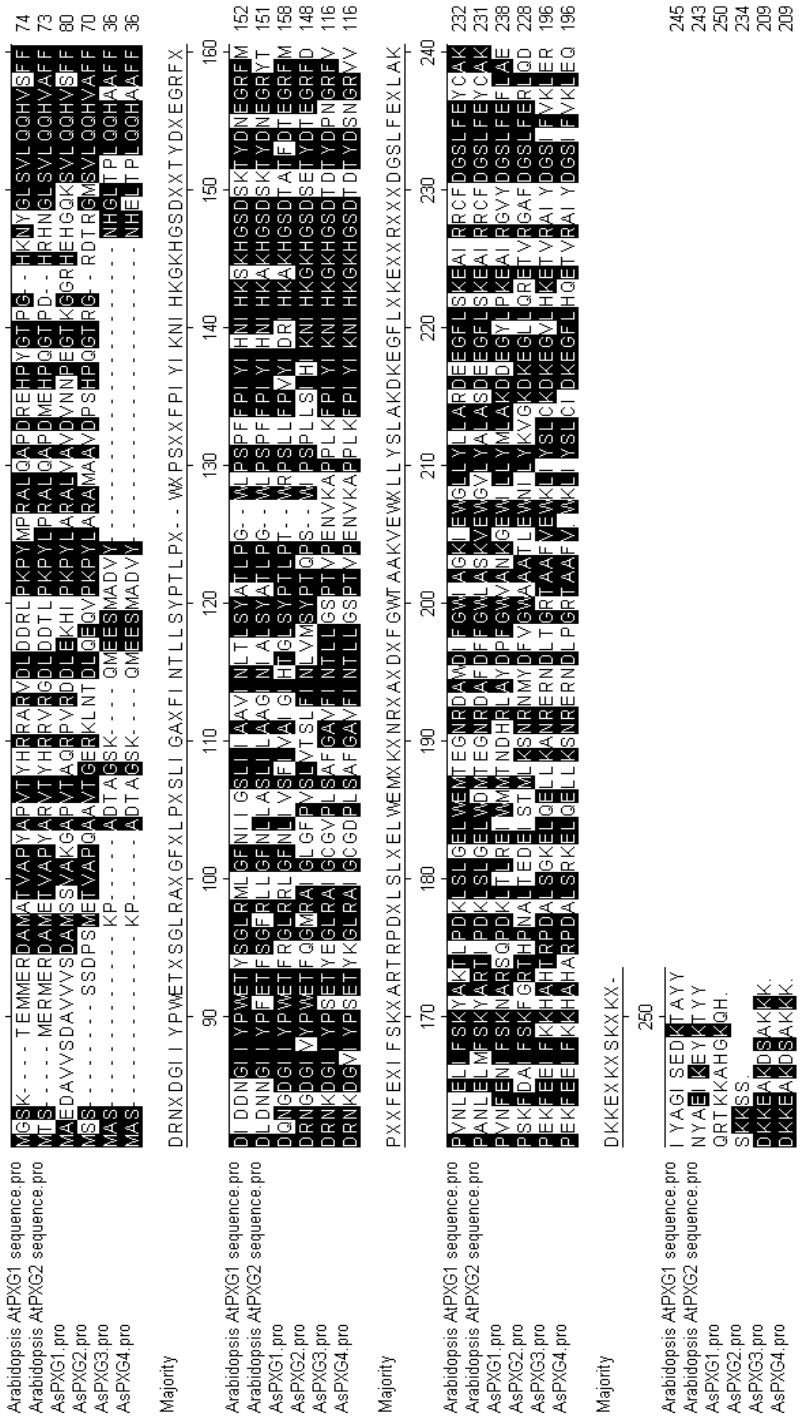


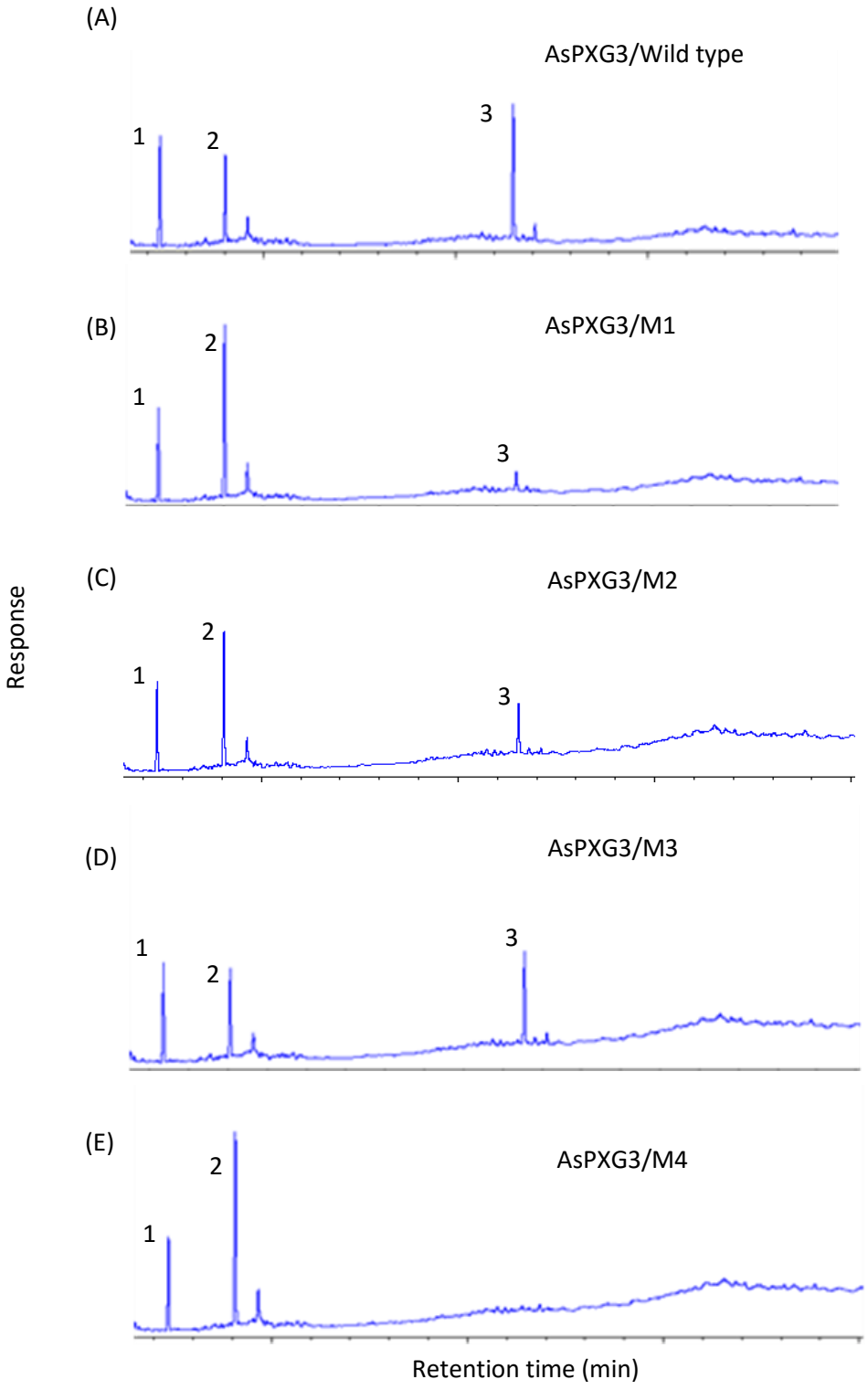
Figure 2: The full length cDNA of *AsPXG4* and the translated open reading frame (ORF)

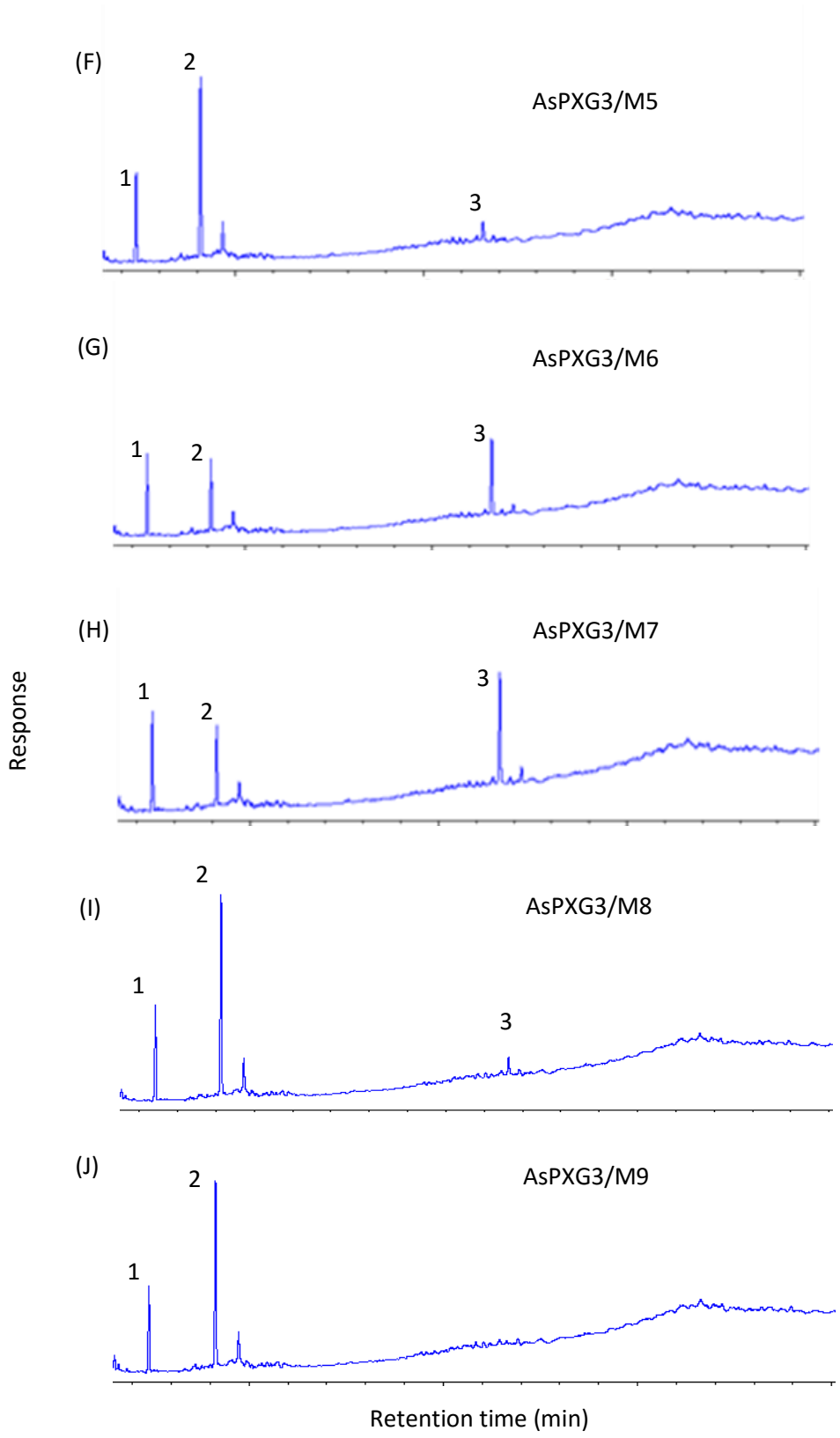
```
ATGGCCTCGAAGCCCGCGGACACTGCAGGGAGCAAGCAGATGGAGGAGTC
 M  A  S  K  P  A  D  T  A  G  S  K  Q  M  E  E  S

CATGGCGGACGTGTACAACCACGaGCTGACGCCGCTGCAGCAGCACGCCG
   M  A  D  V  Y  N  H  E  L  T  P  L  Q  Q  H  A

CCTTCTTCGACCGGAACAAGGACGGCgTCATCTACCCCTCCGAGACCTAC
A  F  F  D  R  N  K  D  G  V  I  Y  P  S  E  T  Y

aAAGGGTTACGCGCGATCGGCTGCGGAGaCCCGCTGTCTGCCTTCGGCGC
 K  G  L  R  A  I  G  C  G  D  P  L  S  A  F  G  A

CGTCTTCATCAACACCTTGCTCGGgTCCCCGACCGTACCGGAGAACGTGA
   V  F  I  N  T  L  L  G  S  P  T  V  P  E  N  V

AGGCTCCGCCTTTGAAATTCCCCATTTACATAAAAAACATCCACAAAGGC
K  A  P  P  L  K  F  P  I  Y  I  K  N  I  H  K  G

AAGCATGGGAGTGATACAGACACGTACGATtCCAATGGAAGGgTTGTTCC
 K  H  G  S  D  T  D  T  Y  D  S  N  G  R  V  V  P

TGAAAAGTTTGAGGAGATATTCAAGAAGCATGCCCACgCAAGaCCTGATG
   E  K  F  E  E  I  F  K  K  H  A  H  A  R  P  D

CCCTATCAcGCAAAGAGCTaCAGGAGTTGCTTAAAtCAAATAGGGAGCGG
A  L  S  R  K  E  L  Q  E  L  L  K  S  N  R  E  R

AAcGATCTAcCaGGACGGACGGCTGCCTTCGTAgAGTGGAAACTTATCTA
 N  D  L  P  G  R  T  A  A  F  V  E  W  K  L  I  Y

CTCATTGTGCAtAGACAAGGAGGGatTTCTTCACcAGGAGACTGTGAGGG
   S  L  C  I  D  K  E  G  F  L  H  Q  E  T  V  R

CAATCTATGATGGtAGCATATTTGTGAAGTTAGAGCaAGACAAGAAGGAA
A  I  Y  D  G  S  I  F  V  K  L  E  Q  D  K  K  E

GCTAAGGATTCTGCCAAGAAGAAATGA
 A  K  D  S  A  K  K  K  .
```

Figure 3: Full length amino acid sequence alignment of *Arabidopsis* peroxygenses (AtPXG1, AtPXG2) and oat peroxygenases (AsPXG1, AsPXG2, AsPXG3 and AsPXG4). The shaded regions indicate amino acids that are identical to each other. Broken lines in the sequence represent gaps introduced for best alignment.



Mutagenesis study




(F)
AsPXG3/M5
1
2
3
(G)
AsPXG3/M6
1
2
3
(H)
AsPXG3/M7
1
2
3
(I)
AsPXG3/M8
1
2
3
(J)
AsPXG3/M9
1
2
Response
Retention time (min)

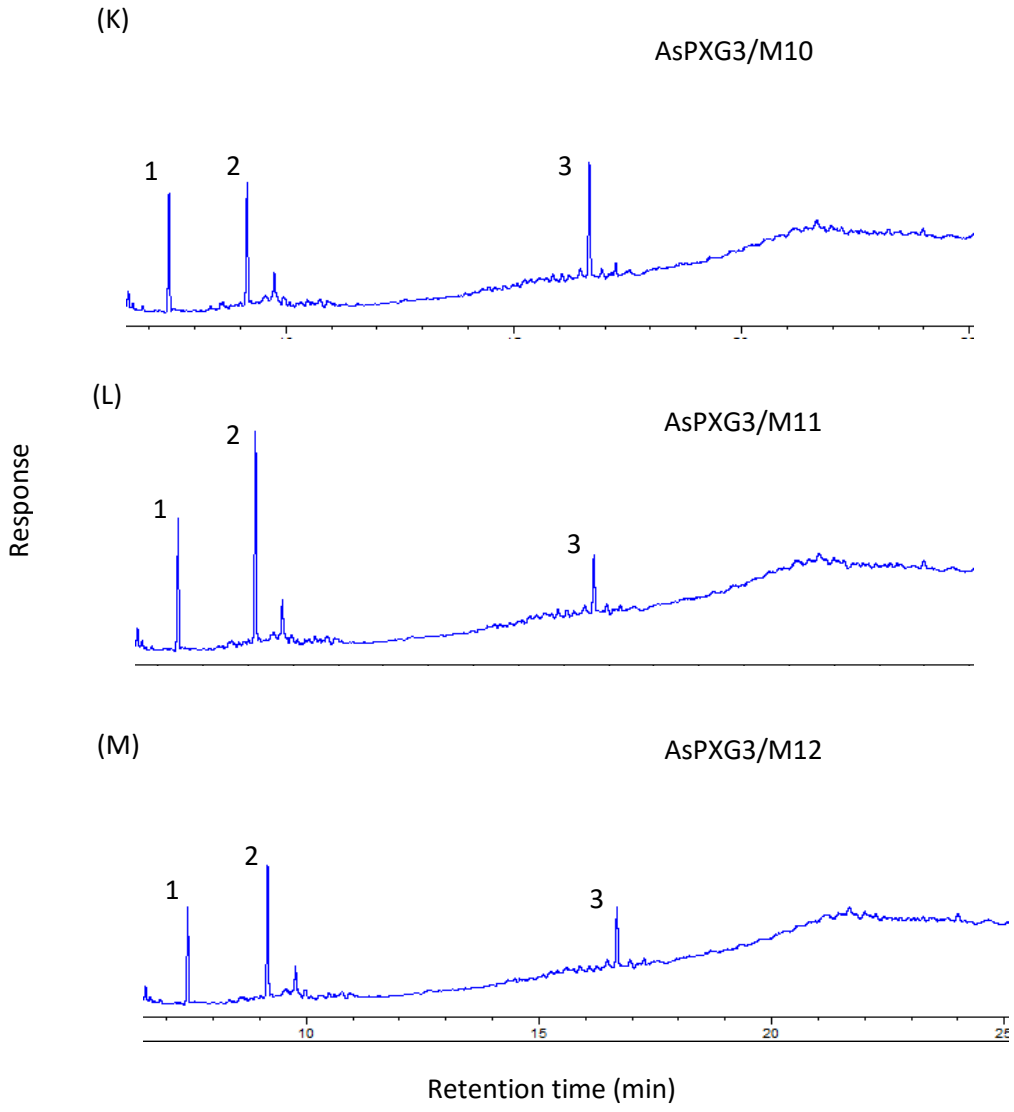


Figure 4: Representative gas chromatographs of activity assays of oat peroxygenase AsPXG3 mutants in the *E. coli* strain in the presence of cumene hydroperoxide as an oxygen donor and oleic acid methyl ester as a substrate. Panel A is the fatty acid products of the wild type, B through M represent the fatty acid products from M1 to M12 mutant respectively. In each panel, fatty acid peaks denoted by numbers are (1) 17:0, (2) 18:1-9*c*, (3) 18:0-9,10 epoxy.

Table 1. Specific activity of AsPXG3/wild type and the three mutants M1, M5 and M8 on different fatty acids.

Substrate	Specific activity (μmol/min/mg)			
	Wild Type	M1	M5	M8
C14:1	167 ± 16.6	68 ± 2.3	47 ± 0.4	56 ± 1.6
C16:1	240 ± 4.3	85 ± 3.2	82 ± 7.2	91 ± 3.1
C18:1	208 ± 2.9	45 ± 3.2	27 ± 3.8	39 ± 1.2
C20:1	203 ± 5.3	62 ± 11.0	52 ± 0.9	34 ± 3.3
C22:1	143 ± 11.3	12 ± 1.5	24 ± 2.6	17 ± 2.2
C24:1	144 ± 6.7	16 ± 5.0	25 ± 1.0	23 ± 15.0
C18:2	685 ± 17.2	258 ± 28.8	279 ± 5.4	246 ± 14.5
C18:3	746 ± 22.0	326 ± 32.7	627 ± 25.7	833 ± 18.3

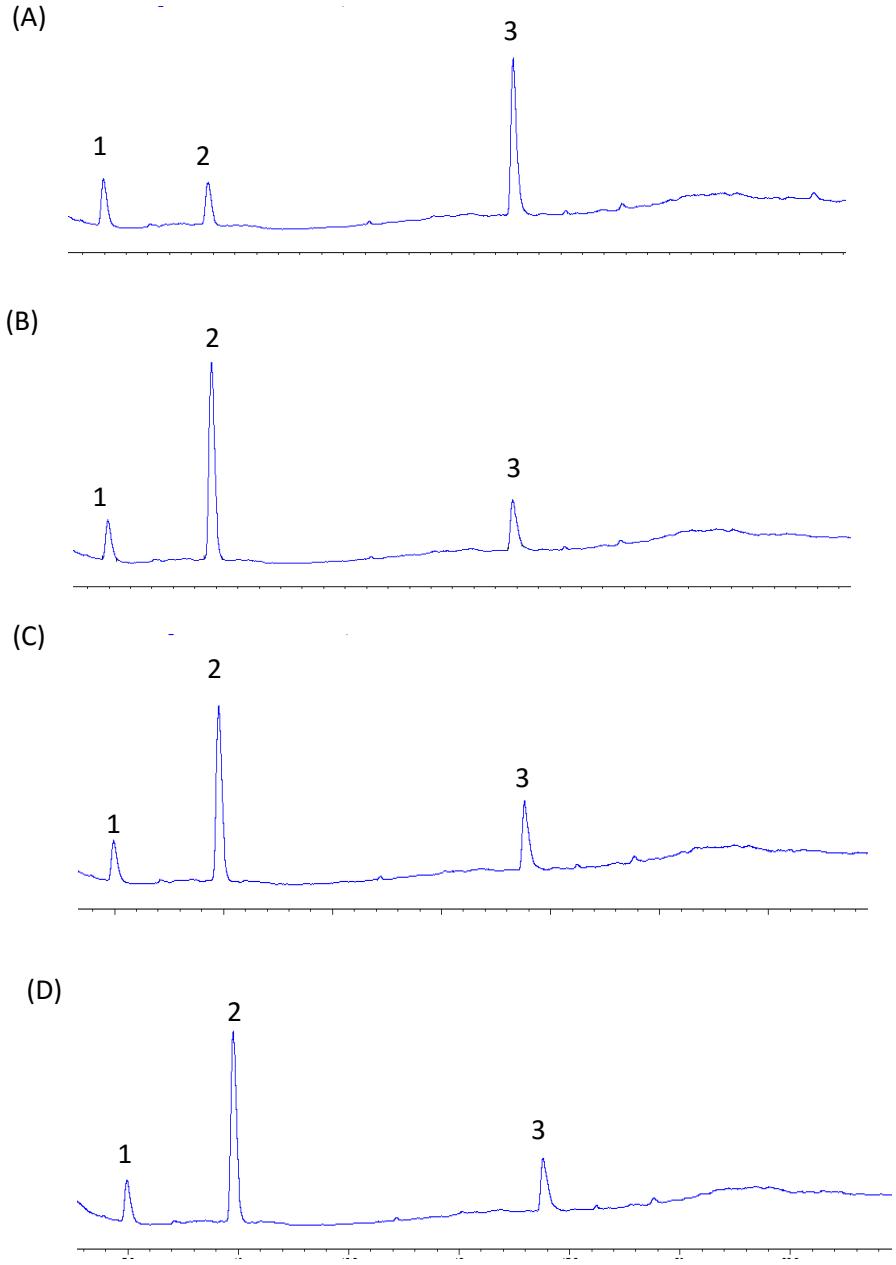


Figure 5. Representative gas chromatographs of activity assays of oat peroxygenase AsPXG3/wild type and three mutants in the *E. coli* strain in the presence of cumene hydroperoxide as an oxygen donor and C18:2-9c,12c as a substrate. Panel A is the fatty acid products of the wild type, B through D represent the fatty acid products M1, M5 and M8 mutant respectively. In each panel, fatty acid peaks denoted by numbers are (1) 17:0, (2) 18:2-9c,12c (3) monoepoxide: 9,10-epoxy-12-octadecenoic acid (most likely according to Blee et al., 2012).

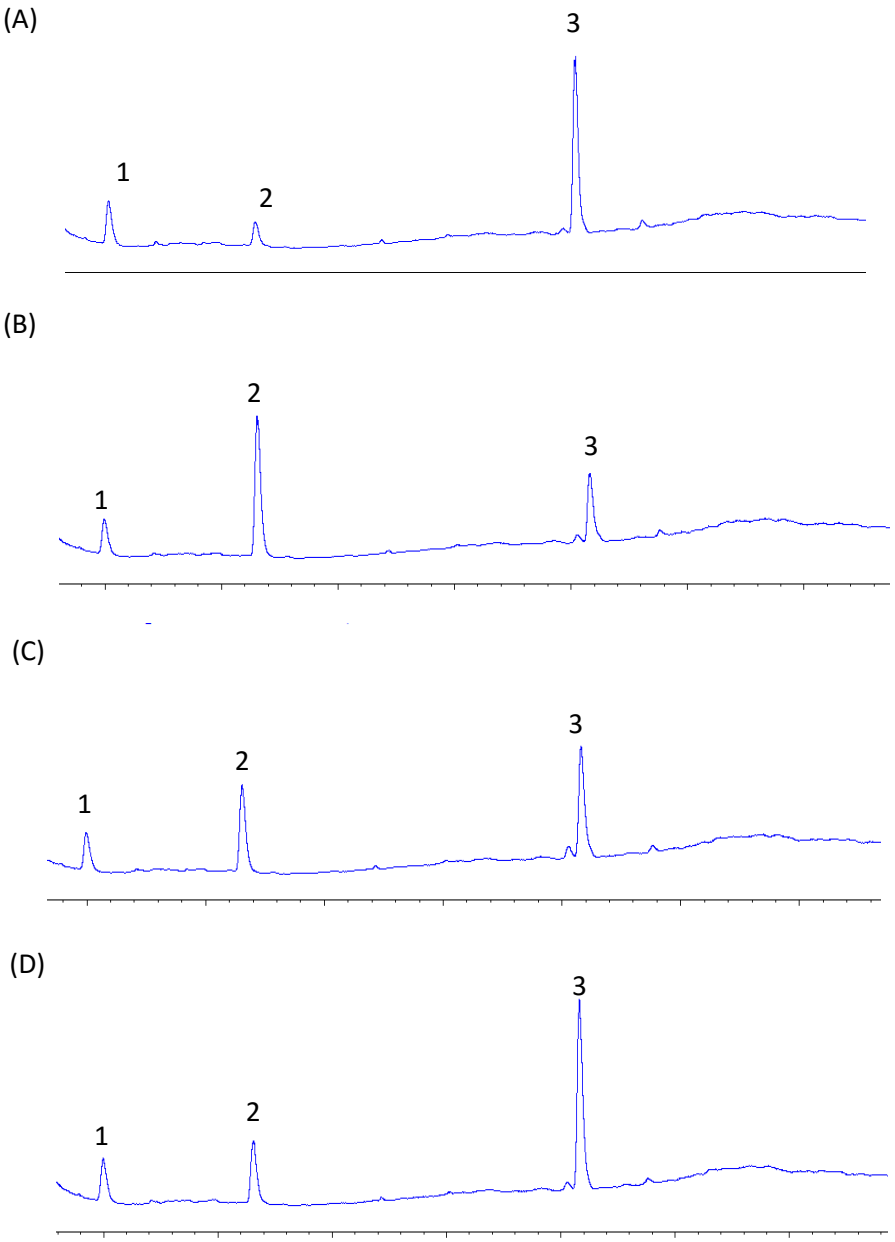


Figure 6. Representative gas chromatographs of activity assays of oat peroxygenase AsPXG3/wild type and three mutants in the *E. coli* strain in presence of cumene hydroperoxide as oxygen donor and C18:3-9c,12c,15c as a substrate. Panel A is the fatty acid products of the wild type, B through D represent the fatty acid products M1, M5 and M8 mutant respectively. In each panel, fatty acid peaks denoted by numbers are (1) 17:0, (2) 18:3-9c,12c,15c (3) diepoxide: 9,10-15,16-diepoxy-12-octadecenoic acid (most likely according to Blee et al., 2012).